

Badura · Ducki · Meyer · Baumgardt  
Schröder *Hrsg.*

# Fehlzeiten-Report 2025

KI und Gesundheit – Möglichkeiten  
nutzen, Risiken bewältigen, Orientierung  
geben

## **Fehlzeiten-Report**

### *Reihenherausgeber*

Bernhard Badura, Fakultät für Gesundheitswissenschaften,  
Universität Bielefeld, Bielefeld, Deutschland

Antje Ducki, Berliner Hochschule für Technik (BHT), Berlin,  
Deutschland

Johanna Baumgardt, Wissenschaftliches Institut der AOK, Berlin,  
Deutschland

Markus Meyer, Wissenschaftliches Institut der AOK (WIdO), Berlin,  
Deutschland

Helmut Schröder, Wissenschaftliches Institut der AOK (WIdO),  
Berlin, Deutschland

Bernhard Badura · Antje Ducki · Markus Meyer ·  
Johanna Baumgardt · Helmut Schröder  
*Hrsg.*

# Fehlzeiten-Report 2025

KI und Gesundheit - Möglichkeiten nutzen, Risiken bewältigen,  
Orientierung geben

*Hrsg.*

Prof. Dr. Bernhard Badura  
Fakultät für Gesundheitswissenschaften  
Universität Bielefeld  
Bielefeld, Deutschland

Prof. Dr. Antje Ducki  
Berliner Hochschule für Technik (BHT)  
Berlin, Deutschland

Markus Meyer  
Wissenschaftliches Institut der AOK (WIdO)  
Berlin, Deutschland

Dr. Johanna Baumgardt  
Forschungsbereich Betriebliche  
Gesundheitsförderung und Heilmittel  
Wissenschaftliches Institut der AOK (WIdO)  
Berlin, Deutschland

Helmut Schröder  
Wissenschaftliches Institut der AOK (WIdO)  
Berlin, Deutschland

#### Fehlzeiten-Report

ISBN 978-3-662-71884-1

ISBN 978-3-662-71885-8 (eBook)

<https://doi.org/10.1007/978-3-662-71885-8>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2025

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jede Person benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des/der jeweiligen Zeicheninhaber\*in sind zu beachten.

Der Verlag, die Autor\*innen und die Herausgeber\*innen gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autor\*innen oder die Herausgeber\*innen übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung: Dr. Christine Lerche

Fotonachweis Umschlag: © istock.com/useng

Springer ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Wenn Sie dieses Produkt entsorgen, geben Sie das Papier bitte zum Recycling.

## Vorwort

---

Künstliche Intelligenz (KI) ist zu einem *der* zentralen Themen unserer Zeit geworden. Sie hat mit atemberaubender Geschwindigkeit Einzug in die Arbeitswelt gehalten, in der ihre Einsatzmöglichkeiten nahezu unbegrenzt sind und stetig erweitert werden. Seit der als „Geburtsstunde der KI“ bezeichneten Konferenz am Dartmouth College, New Hampshire, in den USA im Jahr 1956 bis zum heutigen Tag steigt ihre Relevanz in dem Maß, wie Technologien und Algorithmen fortschreiten und immer komplexere digitale Problemlösungen ermöglichen. Der Mehrwert von KI in der Arbeitswelt äußert sich v. a. in ihrer Fähigkeit, Prozesse zu optimieren und Entscheidungen zu unterstützen, große Datens Mengen effizient zu verarbeiten und daraus Erkenntnisse zu ziehen, die für Menschen allein nicht möglich wären. Damit wird sie zu einem unverzichtbaren Werkzeug in der Arbeitswelt, dessen Bedeutung nicht allein in seiner technologischen Leistungsfähigkeit, sondern auch in seinem Potenzial, Innovationskraft und Wettbewerbsfähigkeit zu steigen, liegt. Gleichzeitig kann die Implementierung von KI in der Arbeitswelt erhebliche Herausforderungen und Gefahren mit sich bringen. Hierzu zählen bspw. Diskriminierung, die Verstärkung sozialer und wirtschaftlicher Ungleichheiten, Arbeitsplatzverluste, eine Beeinträchtigung bzw. Gefährdung der Privatsphäre und die Gefahr der Manipulation, wie bspw. die absichtliche Verbreitung von Falschinformationen mittels KI. Um die Effekte der Einführung von KI in der Arbeitswelt im Allgemeinen und auf die Gesundheit von Beschäftigten im Speziellen vollumfänglich überblicken zu können, bedarf es deshalb unabhängiger, longitudinaler, Setting-spezifischer wissenschaftlicher Studien. Hier zeigt sich hier aktuell eine große Lücke, die zukünftige Forschungsarbeiten systematisch und interdisziplinär schließen sollten.

Vor dem skizzierten Hintergrund beleuchtet der Fehlzeiten-Report 2025 den aktuellen Wissensstand zu Chancen und Risiken beim Einsatz von KI in der Arbeitswelt mit dem Fokus auf die Gesundheit von Beschäftigten. Für einen orientierenden Überblick zum komplexen Zusammenhang von KI und Gesundheit beinhaltet er einführende Beiträge zur technologischen Entstehung, zu ethischen Aspekten, zum Datenschutz sowie zum internationalen Forschungsstand zu KI in der Betrieblichen Gesundheitsförderung und Prävention. Des Weiteren werden in einer Vielzahl wissenschaftlicher Beiträge konkrete Einsatzmöglichkeiten von KI in der Arbeitswelt aus der Perspektive der Gesundheitsförderung dargelegt. Diese erstrecken sich vom Arbeitsschutz und der Gefährdungsbeurteilung über Weiterbildung und Arbeitsqualität bis zu Personalmanagement und Teilhabeförderung.

Zusätzlich zum Schwerpunktthema berichtet der vorliegende Fehlzeiten-Report wie gewohnt über die aktuelle Entwicklung des Krankenstands von Beschäftigten in Deutschland. Detaillierte Analysen von Abrechnungsdaten der 15 Mio. berufstätigen AOK-Versicherten, die im Jahr 2024 in mehr als 1,6 Mio. Betrieben tätig waren, ermöglichen dabei einen umfassenden Überblick über das aktuelle Krankheitsgeschehen. Dabei werden nicht nur soziodemographische und regionale Kriterien, sondern auch die unterschiedlichen Branchen und Berufe differenziert betrachtet. Der Fehlzeiten-Report 2025 berichtet zudem wie gewohnt über die Entwicklung der Krankengeldfälle und -ausgaben von AOK-Mitgliedern sowie über die Krankenstandsentwicklung in der deutschen Bundesverwaltung. Zusätzliche Auswertungen wie etwa zu Arbeitsunfällen, Langzeitarbeitsunfähigkeit, Burnout und Kinderkrankengeld sind zum festen Bestandteil

der Berichterstattung zum Arbeitsunfähigkeitsgeschehen geworden. Vor dem Hintergrund der aktuellen Diskussion um hohe berufliche Fehlzeiten wurde der Report darüber hinaus in diesem Jahr erstmals um einen Beitrag zur Einführung von Karenztagen und zu möglichen Effekten einer Absenkung der Lohnersatzrate erweitert.

Wir hoffen, dass es uns auch in dieser 27. Ausgabe des Fehlzeiten-Reports gelungen ist, Denkanstöße für aktuelle Fragestellungen und Herausforderungen rund um das Thema Arbeit und Gesundheit zu geben. Wir blicken auf eine lange, erfolgreiche Buchreihe zurück, die aufzeigen konnte, mit welchen besonderen Herausforderungen Betriebe, Führungskräfte und Beschäftigte im Themenfeld Arbeit und Gesundheit konfrontiert sind und welche Lösungen des Betrieblichen Gesundheitsmanagements existieren, mit denen die Gesundheit von Beschäftigten gestärkt werden kann.

Der vorliegende Fehlzeiten-Report wäre auch in diesem Jahr nicht ohne die professionelle Zusammenarbeit mit zahlreichen Menschen entstanden. So gilt unser herzlicher Dank allen Autorinnen und Autoren, die ihre fachliche Expertise in den einzelnen Beiträgen in hervorragender Weise eingebracht haben.

Des Weiteren danken wir allen Kolleginnen und Kollegen im Wissenschaftlichen Institut der AOK (WIdO), die an der Erstellung des diesjährigen Fehlzeiten-Reports in vielfältiger Art und Weise beteiligt waren. Zu nennen sind hier vor allem Antonia Düring, Hannes Klawisch und Miriam-Maleika Höltgen, die uns bei der Organisation, Redaktion und Umsetzung dieses Fehlzeiten-Reports exzellent unterstützt haben. Unser Dank gilt weiterhin Susanne Sollmann für das wie immer ausgezeichnete Lektorat. Danken möchten wir gleichermaßen Anja Füssel, Matthias Hell, Melanie Hoburg und Anja Michels im Backoffice des WIdO, ohne deren Unterstützung diese Buchpublikation nicht möglich gewesen wäre. Zudem bedanken wir uns sehr herzlich bei Frau Karg und Frau Dr. Lerche vom Springer-Verlag, die uns hervorragend betreut haben.

Berlin und Bielefeld  
im Juli 2025

# Inhaltsverzeichnis

---

## I Einführung in das Thema

<b>1 Künstliche Intelligenz – eine Schlüsseltechnologie .....</b>	<b>3</b>
<i>Hartmut Hirsch-Kreinsen</i>	
<b>1.1 Der KI-Boom .....</b>	<b>4</b>
<b>1.2 Zur Dynamik der KI .....</b>	<b>5</b>
<b>1.2.1 KI, ein Sammelbegriff .....</b>	<b>5</b>
<b>1.2.2 Weitreichende Versprechungen .....</b>	<b>6</b>
<b>1.3 Diffusion in Grenzen .....</b>	<b>7</b>
<b>1.3.1 Zurückhaltende Verbreitung .....</b>	<b>7</b>
<b>1.3.2 Übersehene Anwendungsprobleme .....</b>	<b>8</b>
<b>1.4 Moderate Folgen für Arbeit .....</b>	<b>10</b>
<b>1.5 Perspektiven .....</b>	<b>11</b>
<b>Literatur .....</b>	<b>12</b>
<b>2 Gesünder, produktiver, automatisierter? Ethische Aspekte von KI im Betrieblichen Gesundheitsmanagement .....</b>	<b>15</b>
<i>Lou Therese Brandner</i>	
<b>2.1 Einleitung: KI zwischen Dystopie und Allheilmittel .....</b>	<b>16</b>
<b>2.2 KI-Ethik als Forschungsfeld .....</b>	<b>16</b>
<b>2.3 Fairness und Antidiskriminierung: Ist KI objektiv? .....</b>	<b>17</b>
<b>2.4 Privatheit versus Transparenz: Werte im Konflikt? .....</b>	<b>19</b>
<b>2.5 Menschliche Autonomie – Aufsicht und Letztentscheidung .....</b>	<b>21</b>
<b>2.6 Fazit .....</b>	<b>23</b>
<b>Literatur .....</b>	<b>23</b>
<b>3 Datenschutz- und KI-rechtliche Aspekte beim Einsatz Künstlicher Intelligenz in Unternehmen .....</b>	<b>27</b>
<i>Rolf Schwartmann und Moritz Köhler</i>	
<b>3.1 KI und Gesundheitsdaten .....</b>	<b>28</b>
<b>3.1.1 Daten im Lebenszyklus eines KI-Systems .....</b>	<b>28</b>
<b>3.1.2 Gesundheitsdaten und der European Health Data Space .....</b>	<b>29</b>
<b>3.2 Datenschutzrechtliche Grundlagen des KI-Einsatzes .....</b>	<b>30</b>
<b>3.2.1 Anwendbarkeit und Verantwortlichkeit .....</b>	<b>30</b>
<b>3.2.2 Rechtmäßigkeit der Verarbeitung .....</b>	<b>31</b>
<b>3.2.3 Verbot der automatisierten Entscheidung im Einzelfall .....</b>	<b>32</b>
<b>3.3 Die Verordnung über künstliche Intelligenz .....</b>	<b>33</b>
<b>3.3.1 Regelungsgegenstand und Akteure der KI-Verordnung .....</b>	<b>33</b>
<b>3.3.2 KI-Kompetenz .....</b>	<b>34</b>
<b>3.3.3 Risikobasierter Ansatz .....</b>	<b>35</b>
<b>3.3.4 Wechsel in die Anbieterrolle .....</b>	<b>36</b>
<b>3.4 Zusammenfassung und Ausblick .....</b>	<b>36</b>
<b>Literatur .....</b>	<b>36</b>

4	<b>„KI als Medizin“ für das Gesundheitswesen?</b>	39
	<i>Martin Roesler</i>	
4.1	<b>Handlungsbedarf und veränderte Rahmenbedingungen für die Medizin</b>	41
4.2	<b>Medizin ohne medizinisches Personal?</b>	42
4.3	<b>Entlastung von Routine- und Dokumentationsaufgaben</b>	43
4.4	<b>Konzentration auf Kernkompetenzen – mehr Zeit für Patienten und lebenslanges Lernen</b>	44
4.5	<b>Höhere Sicherheit bei der Entscheidungsfindung – psychische Entlastung für Ärztinnen und Ärzte</b>	44
4.6	<b>Risiken für medizinisches Personal – Überwachung, Leistungsdruck und Statusverlust</b>	45
4.7	<b>Gesundheitsberufe mutmaßlich unterschiedlich von Automatisierung betroffen</b>	46
4.8	<b>Die allgemeine Transformation des Arbeitsmarktes</b>	47
4.9	<b>Fazit – Chancen und Fallstricke für Gesundheitsfachkräfte</b>	47
	<b>Literatur</b>	49
5	<b>Von der Personalplanung bis zum Gesundheitsmanagement: Nutzungsmöglichkeiten von künstlicher Intelligenz im Personalwesen</b>	51
	<i>Alexander Dregger</i>	
5.1	<b>Warum sich ein Einsatz von KI im HR lohnt</b>	52
5.2	<b>Der Personallebenszyklus und mögliche Einsatzgebiete von KI</b>	53
5.2.1	<b>Personalplanung</b>	54
5.2.2	<b>Personalmarketing</b>	55
5.2.3	<b>Personalauswahl</b>	55
5.2.4	<b>Onboarding von Personal</b>	58
5.2.5	<b>Personalentwicklung</b>	59
5.2.6	<b>Personalmanagement und Personalbindung</b>	60
5.3	<b>Fallstricke bei der Implementierung von KI im HR-Bereich und mögliche Lösungen</b>	61
5.4	<b>Fazit</b>	62
	<b>Literatur</b>	62
6	<b>Künstliche Intelligenz in der betrieblichen Prävention und Gesundheitsförderung</b>	65
	<i>David Matusiewicz</i>	
6.1	<b>Einleitung</b>	67
6.2	<b>KI und das Wohlbefinden der Beschäftigten</b>	67
6.2.1	<b>Personalisierte Gesundheitsförderung am Arbeitsplatz</b>	68
6.2.2	<b>Stressreduktion durch KI-gestützte Systeme</b>	68
6.2.3	<b>Psychische Gesundheit und KI</b>	69
6.2.4	<b>Digitale Gesundheitsanwendungen in der Betrieblichen Gesundheitsförderung und KI</b>	70
6.2.5	<b>Ergonomische Gesundheitsförderung am Arbeitsplatz durch KI</b>	70
6.3	<b>KI und Produktivität</b>	71
6.3.1	<b>Automatisierung von Routineaufgaben</b>	71
6.3.2	<b>Assistenzsysteme und kollaborative Roboter</b>	71
6.3.3	<b>KI-gestützte Analyse und Optimierung von Arbeitsprozessen</b>	72
6.3.4	<b>Verbesserung der Mensch-Maschine-Kollaboration</b>	72
6.4	<b>KI als Chance für gesunde Arbeitsbedingungen</b>	72
6.4.1	<b>Wearables zur Gesundheitsüberwachung</b>	72

## Inhaltsverzeichnis

6.4.2	Adaptive Assistenzsysteme in der Produktion .....	73
6.4.3	Robotik für ergonomische Entlastung.....	74
6.4.4	KI in der Gefahrenvermeidung am Arbeitsplatz .....	75
6.4.5	Einsatz von KI zur Unfallverhütung .....	75
6.5	<b>KI und Gesundheitskompetenz.....</b>	75
6.5.1	KI-gestützte Schulungs- und Lernplattformen .....	76
6.5.2	KI-basierte personalisierte Trainingsprogramme .....	77
6.6	<b>Herausforderungen und ethische Aspekte .....</b>	77
6.6.1	Datenschutz und KI in der Arbeitswelt .....	77
6.6.2	Mensch-Maschine-Interaktion und Akzeptanz .....	77
6.6.3	Ethische Fragen beim Einsatz von KI .....	78
6.6.4	Regulierungen und rechtliche Rahmenbedingungen .....	78
6.7	<b>Fazit und Ausblick .....</b>	79
	Literatur .....	80
7	<b>KI in der Betrieblichen Gesundheitsförderung und Prävention – ausgewählte Perspektiven zum internationalen Forschungsstand .....</b>	83
	<i>Martin Lange und Andrea Schaller</i>	
7.1	<b>KI: Grundlagen und Einsatzmöglichkeiten in der Gesundheitsversorgung .....</b>	85
7.2	<b>Herausforderungen der Evidenzentwicklung und -bewertung von KI-Anwendungen .....</b>	86
7.3	<b>Aktueller Stand zu KI in der BGF .....</b>	87
7.4	<b>BGF- und Präventions-bezogene Einsatzbereiche der identifizierten KI-Anwendungen .....</b>	88
7.5	<b>Identifizierte KI-Anwendungen und Technologien .....</b>	98
7.6	<b>Evidenz der KI-Anwendungen .....</b>	98
7.7	<b>Diskussion und Implikationen .....</b>	99
7.8	<b>Ausblick und Fazit .....</b>	103
	Literatur .....	103
<b>II</b>	<b>Gesundheitsorientierter Einsatz von KI – Ebene der Organisation</b>	
8	<b>Auswirkungen der Künstlichen Intelligenz auf Arbeit, Beschäftigung und Gesundheit .....</b>	109
	<i>Martin Braun</i>	
8.1	<b>Einleitung .....</b>	111
8.2	<b>Kernkonzepte der Arbeitsgestaltung .....</b>	111
8.2.1	<b>Verständnis und Grundlagen .....</b>	111
8.2.2	<b>Produktivitätssteigerung durch rationale Arbeitsteilung .....</b>	112
8.2.3	<b>Soziale Wirkung der Arbeitsteilung .....</b>	112
8.2.4	<b>Sozio-technisches Arbeitssystem .....</b>	113
8.2.5	<b>Humanisierung von Arbeit .....</b>	113
8.2.6	<b>Informatisierung von Arbeit .....</b>	113
8.3	<b>Künstliche Intelligenz als Mittel der Informatisierung .....</b>	114
8.3.1	<b>Verständnis und Historie .....</b>	114
8.3.2	<b>Anwendungen der Künstlichen Intelligenz .....</b>	114
8.3.3	<b>Grenzen der KI-Anwendung .....</b>	115
8.3.4	<b>Ethische Implikationen .....</b>	115

8.4	<b>Auswirkungen des KI-Einsatzes auf die Arbeit</b> .....	115
8.4.1	Ebene der Arbeitsmärkte und Berufsbilder .....	116
8.4.2	Ebene der Arbeitsstrukturen und Tätigkeiten .....	117
8.4.3	Ebene der Qualifikationen .....	118
8.5	<b>Gesundheitliche Implikationen des KI-Einsatzes</b> .....	119
8.5.1	Stufen der Technisierung von Arbeit .....	119
8.5.2	Gesundheit als autoregulativer Adoptionsprozess .....	121
8.5.3	Komplementarität von Humanisierung und Rationalisierung .....	121
8.5.4	Menschengerechte Arbeitsgestaltung .....	122
8.6	<b>Fazit</b> .....	122
	Literatur .....	123
9	<b>Künstliche Intelligenz als Gegenstand und Instrument des Arbeitsschutzes</b> .....	125
	<i>Lars Adolph</i>	
9.1	<b>Einleitung</b> .....	127
9.2	<b>Die EU-KI-Verordnung</b> .....	129
9.2.1	KI im Sinne der EU-KI-Regulation – Versuche einer Definition .....	129
9.2.2	Gruppierung von KI-Systemen nach Risikokriterien der EU-KI-Verordnung .....	130
9.2.3	Anforderungen an KI-Systeme nach EU-KI-Verordnung – Konkretisierung durch Normung .....	131
9.2.4	Menschliche Aufsicht: Lösung oder Problem? .....	132
9.3	<b>Die Sicht des Arbeitsschutzes: KI-Systeme als sichere Arbeitsmittel gestalten und im Betrieb sicher anwenden</b> .....	133
9.3.1	KI-Systeme als Gegenstand von Gefährdungsbeurteilungen .....	134
9.3.2	Besonderheit der Gefährdungsbeurteilung für Systeme des Algorithmischen Managements .....	135
9.4	<b>Chancen von KI für den Arbeitsschutz und die menschengerechte Arbeitsgestaltung</b> .....	135
9.4.1	KI als Unterstützungsinstrument für Gefährdungsbeurteilungen .....	135
9.4.2	Sensorik und KI-Analysen für Gefährdungsbeurteilungen .....	136
9.4.3	Smarte Arbeitsschutzzvorschriften .....	137
9.5	<b>Fazit: Mit KI-Technologien sichere, gesunde und menschengerechte Arbeit gestalten</b> .....	137
	Literatur .....	138
10	<b>Personalmix und Künstliche Intelligenz: Erfolgsfaktoren, Herausforderungen und Handlungsempfehlungen</b> .....	141
	<i>Sascha Stowasser</i>	
10.1	<b>Vielfaltbewusste Organisationen</b> .....	143
10.1.1	Grundlagen sowie Vor- und Nachteile .....	143
10.1.2	Diversität als Wettbewerbsvorteil .....	143
10.1.3	Diversität strategisch fördern .....	144
10.2	<b>Integration von KI in vielfältigen Teams</b> .....	144
10.3	<b>Herausforderungen bei der Integration von KI in vielfältigen Teams</b> .....	145
10.3.1	Technologische und soziale Herausforderungen der KI-Integration in diversen Teams .....	145
10.3.2	Akzeptanzprobleme und Weiterbildungsbedarf in KI-gestützten Teams .....	146
10.3.3	Veränderungen in der Teamdynamik durch KI in diversen Teams .....	146
10.3.4	Herausforderungen in der Kommunikation zwischen Mensch und KI .....	146
10.3.5	Verantwortung und ethische Fragen in diversen Teams mit KI .....	147

<b>10.4 Handlungsempfehlungen für Organisationen: Erfolgreiche Integration von KI in vielfältige Teams .....</b>	148
<b>10.4.1 Akzeptanz und geeignete Organisationskultur schaffen .....</b>	148
<b>10.4.2 Kompetenzen für eine KI-gestützte Arbeitswelt stärken .....</b>	149
<b>10.4.3 Rollen- und Verantwortungsstrukturen definieren .....</b>	149
<b>10.4.4 Ethische Richtlinien und faire KI-Nutzung fördern .....</b>	150
<b>10.4.5 Regelmäßige Evaluationsprozesse etablieren .....</b>	150
<b>10.4.6 Vielfalt und Technologie als Einheit denken .....</b>	150
<b>10.5 Beispiele für Good Practice .....</b>	151
<b>10.6 Fazit: KI als Chance für Vielfalt und nachhaltige Arbeitsgestaltung .....</b>	151
<b>Literatur .....</b>	152
<b>11 Möglichkeiten und Bewertung von digitalem Betrieblichem Gesundheitsmanagement (BGM) mit Künstlicher Intelligenz (KI) .....</b>	155
<i>Carolin Wienrich und Susanne Völter-Mahlknecht</i>	
<b>11.1 Einleitung .....</b>	156
<b>11.2 Potenziale für die Nutzung von KI im Bereich BGM mit besonderem Fokus auf BGF .....</b>	157
<b>11.3 Grenzen und Herausforderungen für die Nutzung von KI im Bereich BGM mit besonderem Fokus auf BGF .....</b>	159
<b>11.4 Einführung von KI im Bereich BGM und BGF .....</b>	160
<b>11.5 KI als „persönlicher Gesundheitsmanager“ .....</b>	162
<b>11.6 Fazit .....</b>	163
<b>11.7 Ausblick .....</b>	164
<b>Literatur .....</b>	164
<b>12 Betriebliche Gesundheitsrisiken erkennen, bevor sie entstehen? Wie KI die Gefährdungsbeurteilung unterstützen kann .....</b>	167
<i>Thomas Lennefer, Moritz Schneider und Ulrike Rösler</i>	
<b>12.1 Einleitung .....</b>	168
<b>12.2 Die Gefährdungsbeurteilung: Ein zentrales Präventionsinstrument im Betrieb .....</b>	169
<b>12.2.1 Beispiel I: Fortschritte in der Sturzprävention am Arbeitsplatz: Herausforderungen und Potenziale durch Künstliche Intelligenz .....</b>	170
<b>12.2.2 Beispiel II: Potenziale von KI für die Prävention von physischer und psychischer Gewalt am Arbeitsplatz .....</b>	173
<b>12.3 Fazit .....</b>	176
<b>Literatur .....</b>	177
<b>13 Mensch-KI-Symbiose: Wie der Einsatz von künstlicher Intelligenz im Unternehmen gelingen kann .....</b>	179
<i>Julia Bosbach und Thomas Lennefer</i>	
<b>13.1 Einleitung .....</b>	180
<b>13.2 Unterschied zwischen künstlicher Intelligenz und menschlicher Intelligenz .....</b>	181
<b>13.3 Menschenzentrierte KI-Integration in Organisationen .....</b>	182
<b>13.4 Methodik .....</b>	183
<b>13.5 Ergebnisse .....</b>	185
<b>13.6 Diskussion .....</b>	189
<b>13.7 Praktische Implikationen .....</b>	190
<b>Literatur .....</b>	191

14	<b>en[AI]ble – ein Weiterbildungsangebot zur mitarbeitergerechten Implementierung und Gestaltung von KI in KMU</b> ..... <i>Sebastian Terstegen und Martina Zahn</i>	193
14.1	<b>Einleitung</b> .....	194
14.2	<b>KI verändert die Arbeitswelt</b> .....	195
14.3	<b>Herausforderungen und Lösungsansätze bei der Einführung von KI in KMU</b> .....	196
14.3.1	Kriterien für die Bewertung der Gestaltung von betrieblichen KI-Anwendungen .....	197
14.4	<b>Erfolgsfaktoren für ein gelingendes Change-Management der KI-Einführung</b> .....	200
14.4.1	Partizipation und Mitbestimmung .....	201
14.4.2	Strukturiertes Vorgehen .....	202
14.5	<b>Fazit und Ausblick</b> .....	203
	Literatur .....	204
15	<b>Künstliche Intelligenz: Neue Potenziale für „gute Arbeit“ im Büro?</b> ..... <i>Tobias Kämpf und Thomas Liihr</i>	207
15.1	<b>Einleitung</b> .....	208
15.2	<b>KI braucht Arbeit</b> .....	209
15.3	<b>Wie KI die Arbeit im Büro verändert</b> .....	212
15.3.1	Neue Potenziale für gute Arbeit .....	212
15.3.2	Entwertung statt „guter Arbeit“: Die Potenziale bleiben ungenutzt .....	214
15.4	<b>Mehr als Technik: KI muss gestaltet werden</b> .....	215
	Literatur .....	217
<b>III</b>	<b>Gesundheitsorientierter Einsatz von KI – Ebene des Individuums</b>	
16	<b>Einführung von KI: Anforderungen an den Transformationsprozess und an KI am Arbeitsplatz aus Sicht der Beschäftigten</b> ..... <i>Sabine Pfeiffer</i>	221
16.1	<b>Theoretischer Hintergrund</b> .....	222
16.2	<b>Methodisches Vorgehen</b> .....	223
16.3	<b>Zentrale Ergebnisse</b> .....	225
16.4	<b>Diskussion und Ausblick</b> .....	232
	Literatur .....	233
17	<b>Künstliche Intelligenz und digitale Selbstvermessung – Technik als Trostersatz in erschöpfenden Zeiten</b> ..... <i>Stefan Selke</i>	235
17.1	<b>Konvergenz digitaler Selbstvermessung und künstlicher Intelligenz</b> .....	236
17.2	<b>Metamorphosen der Mensch-Technik-Beziehung im Kontext KI-basierter Selbstvermessung</b> .....	237
17.2.1	Differenzierung des Rollenverständnisses von KI-Anwendungen .....	238
17.2.2	KI als Werkzeug .....	239
17.2.3	KI als Assistent .....	239
17.2.4	KI als Sozialpartner .....	240
17.2.5	Vom Sozialpartner zum Super-Ego .....	241

17.3	<b>Schleichende Veränderungen des gesellschaftlichen Referenzrahmens für KI-basiertes Self-Tracking</b> .....	241
17.3.1	Kontingenzdilemma in metrischen Kulturen .....	242
17.3.2	Rationale Diskriminierung.....	243
17.4	<b>KI und Selbstvermessung im Kontext von Zukunftsnarrativen</b> .....	243
	Literatur .....	245
18	<b>Arbeitsqualität im Kontext KI-gestützter Arbeitsorganisation</b> .....	247
	<i>Swantje Robelski, Sophie-Charlotte Meyer und Matthias Hartwig</i>	
18.1	<b>Einleitung: Künstliche Intelligenz und die Veränderung der Arbeitsorganisation</b> ...	248
18.1.1	Was bedeutet KI-gestützte Arbeitsorganisation? .....	248
18.1.2	Verbreitung und Einsatz .....	249
18.1.3	Folgen KI-gestützter Arbeitsorganisation für die Arbeitsqualität .....	250
18.2	<b>Daten und Methode</b> .....	250
18.3	<b>Ergebnisse</b> .....	251
18.4	<b>Diskussion</b> .....	254
18.5	<b>Fazit und Ausblick</b> .....	255
	Literatur .....	256

## IV Perspektiven

19	<b>Chancen von KI im Arbeitsschutz aus Sicht der deutschen Unternehmen</b> <i>Susanne Wagenmann und Elisa Clauß</i>	261
19.1	<b>Die Nutzung von KI in deutschen Unternehmen</b> .....	263
19.1.1	KI als Booster für den Wirtschaftsstandort Deutschland .....	263
19.1.2	Branchen und Beschäftigte nutzen KI noch unterschiedlich intensiv .....	263
19.2	<b>Chancen und Risiken von KI für die Sicherheit und Gesundheit der Beschäftigten</b> ..	265
19.2.1	KI kann die Beschäftigten aktiv schützen und negative psychische Belastungsfaktoren reduzieren .....	265
19.2.2	KI kann bei der Umsetzung des Arbeitsschutzes unterstützen .....	266
19.2.3	Risiken von KI für den Arbeitsschutz .....	267
19.3	<b>Die neue KI-Verordnung als europäischer Rahmen für KI-Nutzung</b> .....	268
19.3.1	Arbeitsschutz im Wandel: Neue Anforderungen durch Normung .....	268
19.3.2	Bürokratie als Innovationshemmnis für den Einsatz von KI .....	269
19.3.3	Betriebliche KI-Nutzung durch klare Vorgaben stärken .....	270
19.4	<b>KI-Nutzung in der Praxis</b> .....	270
19.4.1	KI-Nutzung in der Praxis profitiert von einer offiziellen Einführung .....	270
19.4.2	KI erfolgreich in der Praxis einführen durch den „Digital Mentor“ .....	271
19.5	<b>Fazit</b> .....	272
	Literatur .....	273
20	<b>KI für Gute Arbeit nutzbar machen</b> <i>Oliver Suchy</i>	275
20.1	<b>KI – Chancen und Risiken</b> .....	276
20.2	<b>Transparenz und Autonomie</b> .....	277
20.3	<b>KI-Vorgehensmodelle</b> .....	278

20.4	<b>Die europäische KI-Verordnung</b>	279
20.5	<b>Handlungs- und Konkretisierungsbedarf</b>	279
	Literatur	280
21	<b>Künstliche Intelligenz als strategische Führungsaufgabe</b>	283
	<i>Rahild Neuburger</i>	
21.1	<b>Einführung: KI als neues Element in der Arbeitswelt</b>	285
21.2	<b>Künstliche Intelligenz – Ausprägungen, Potenziale und Grenzen</b>	285
21.2.1	<b>Ausprägungen der künstlichen Intelligenz</b>	285
21.2.2	<b>Potenziale und Grenzen des Einsatzes künstlicher Intelligenz in Organisationen</b>	287
21.3	<b>Gesundheitliche Folgen der künstlichen Intelligenz für Beschäftigte</b>	288
21.4	<b>Künstliche Intelligenz für Führungsaufgaben</b>	290
21.4.1	<b>Administrative Führungsaufgaben</b>	290
21.4.2	<b>Strategische und organisationale Führungsaufgaben</b>	291
21.4.3	<b>Personalführung und HR</b>	291
21.4.4	<b>Selbstführung</b>	291
21.5	<b>Gestaltung und Steuerung der Implementierung von KI</b>	292
21.5.1	<b>Planung: Prüfung konkreter Einsatzfelder</b>	292
21.5.2	<b>Einführung von KI</b>	294
21.5.3	<b>Verfestigung der KI-unterstützten Arbeitswelt</b>	296
21.6	<b>Zusammenfassung: Führung mit KI und Führung der KI</b>	297
	Literatur	298
22	<b>Zukunftsperspektiven von KI in der Betrieblichen Gesundheitsförderung</b>	301
	<i>Markus H. Dahn und Christian Jauch</i>	
22.1	<b>Einleitung</b>	303
22.2	<b>KI: Einsatz und Implikationen</b>	305
22.2.1	<b>Veränderung von Arbeitsprozessen und Anforderungen</b>	305
22.2.2	<b>Chancen und Risiken im Kontext von Gesundheit und BGF</b>	305
22.3	<b>Physische gesundheitliche Auswirkungen</b>	307
22.3.1	<b>Veränderungen der Arbeitsbedingungen durch KI</b>	307
22.3.2	<b>Ergonomische Herausforderungen und Gestaltung der Arbeitsplätze</b>	307
22.3.3	<b>Körperliche Beschwerden und Erkrankungen</b>	308
22.3.4	<b>Präventive Maßnahmen und Betriebliche Gesundheitsförderung (BGF)</b>	308
22.4	<b>Psychische Auswirkungen</b>	309
22.4.1	<b>Psychische Belastungen durch den Einsatz von KI</b>	309
22.4.2	<b>Stress und Überlastung durch Informationsflut und Multitasking</b>	309
22.4.3	<b>Emotionale Auswirkungen und digitale Erschöpfung</b>	309
22.4.4	<b>Präventive Maßnahmen und psychische Gesundheitsförderung</b>	310
22.5	<b>Ganzheitliche Betrachtung und Diskussion</b>	311
22.5.1	<b>Zusammenführung der physischen und psychischen Auswirkungen</b>	311
22.5.2	<b>Chancen und Potenziale von KI zur Gesundheitsförderung</b>	311
22.5.3	<b>Herausforderungen und ethische Implikationen</b>	311
22.5.4	<b>Handlungsempfehlungen für Unternehmen und Führungskräfte</b>	312
22.6	<b>Zusammenfassung</b>	313
22.7	<b>Fazit und Ausblick</b>	314
	Literatur	314

## V Daten und Analysen zu krankheitsbedingten Fehlzeiten

<b>23 Wie geht es den Beschäftigten in Deutschland – Ergebnisse aus Befragungen zwischen 2020 und 2025 .....</b>	<b>319</b>
<i>Johanna Baumgardt und Hannes Klawisch</i>	
23.1 Theoretischer Hintergrund und Fragestellungen .....	320
23.2 Methodik .....	321
23.3 Ergebnisse und Diskussion .....	325
23.4 Zusammenfassung und Ableitungen für die betriebliche Praxis .....	338
23.5 Anhang: Klassifikation der Berufe .....	340
Literatur .....	341
<b>24 Krankheitsbedingte Fehlzeiten in der deutschen Wirtschaft im Jahr 2024 .....</b>	<b>345</b>
<i>Antje Schenkel, Markus Meyer, Antonia Düring und Johanna Baumgardt</i>	
24.1 Überblick über die krankheitsbedingten Fehlzeiten im Jahr 2024 .....	347
24.2 Datenbasis und Methodik .....	350
24.3 Allgemeine Krankenstandsentwicklung .....	353
24.4 Verteilung der Arbeitsunfähigkeit .....	356
24.5 Kurz- und Langzeiterkrankungen .....	357
24.6 Krankenstandsentwicklung in den einzelnen Branchen .....	358
24.7 Einfluss der Alters- und Geschlechtsstruktur .....	363
24.8 Fehlzeiten nach Bundesländern .....	366
24.9 Fehlzeiten nach Ausbildungsabschluss und Vertragsart .....	371
24.10 Fehlzeiten nach Berufsgruppen .....	374
24.11 Fehlzeiten nach Wochentagen .....	375
24.12 Arbeitsunfälle .....	377
24.13 Krankheitsarten im Überblick .....	382
24.14 Die häufigsten Einzeldiagnosen .....	388
24.15 Krankheitsarten nach Branchen und Berufen .....	390
24.16 Langzeitfälle nach Krankheitsarten .....	403
24.17 Krankheitsarten nach Diagnoseuntergruppen .....	406
24.18 Burnout-bedingte Fehlzeiten .....	409
24.19 Arbeitsunfähigkeiten nach Städten 2024 .....	412
24.20 Inanspruchnahme von Krankengeld bei Erkrankung des Kindes .....	415
Literatur .....	421
<b>25 Krankheitsbedingte Fehlzeiten nach Branchen im Jahr 2024 .....</b>	<b>423</b>
<i>Antje Schenkel, Antonia Düring und Markus Meyer</i>	
25.1 Banken und Versicherungen .....	424
25.2 Baugewerbe .....	440
25.3 Dienstleistungen .....	458
25.4 Energie, Wasser, Entsorgung und Bergbau .....	478
25.5 Erziehung und Unterricht .....	497
25.6 Gesundheits- und Sozialwesen .....	514
25.7 Handel .....	533
25.8 Land- und Forstwirtschaft .....	551
25.9 Metallindustrie .....	567
25.10 Öffentliche Verwaltung .....	586

25.11	<b>Verarbeitendes Gewerbe</b> .....	603
25.12	<b>Verkehr und Transport</b> .....	626
26	<b>Entwicklung der Krankengeldfälle und -ausgaben bei AOK-Mitgliedern im Jahr 2024</b> .....	643
	<i>Miriam Räker und Reinhard Schwanke</i>	
26.1	<b>Einführung</b> .....	644
26.2	<b>Einordnung der Datenquelle</b> .....	645
26.3	<b>Entwicklung des Krankengelds</b> .....	646
26.3.1	<b>Krankengeldfallzahlen</b> .....	647
26.3.2	<b>Krankengeldfalldauern</b> .....	648
26.3.3	<b>Krankengeldausgaben nach Diagnosen</b> .....	651
26.3.4	<b>Einfluss des Alters</b> .....	652
26.4	<b>Zusammenfassung</b> .....	654
	<b>Literatur</b> .....	656
27	<b>Krankheitsbedingte Fehlzeiten in der Bundesverwaltung und Bilanz von 15 Jahren Betrieblichen Gesundheitsmanagements in der unmittelbaren Bundesverwaltung</b> .....	657
	<i>Annette Schlippahk und Björn Wegener</i>	
27.1	<b>Gesundheitsmanagement in der Bundesverwaltung</b> .....	659
27.2	<b>Bilanz über 15 Jahre Betriebliches Gesundheitsmanagement in der unmittelbaren Bundesverwaltung</b> .....	659
27.2.1	<b>Einführung</b> .....	659
27.2.2	<b>Was hat sich seit 2009 im BGM der unmittelbaren Bundesverwaltung getan?</b> .....	660
27.2.3	<b>Herausforderungen bei der Umsetzung des BGM</b> .....	661
27.2.4	<b>Umsetzungsstand des BGM</b> .....	662
27.2.5	<b>Die Ergebnisse zum Stand des BGM in der unmittelbaren Bundesverwaltung bezogen auf die Prozessschritte im Managementzyklus</b> .....	663
27.2.6	<b>Fazit und Blick in die Zukunft: Weiterentwicklung des BGM</b> .....	664
27.3	<b>Überblick über die krankheitsbedingten Abwesenheitszeiten im Jahr 2023</b> .....	665
27.3.1	<b>Methodik der Datenerfassung</b> .....	665
27.3.2	<b>Allgemeine Entwicklung der Abwesenheitszeiten</b> .....	665
27.3.3	<b>Dauer der Erkrankung</b> .....	665
27.3.4	<b>Abwesenheitstage nach Laufbahngruppen</b> .....	667
27.3.5	<b>Abwesenheitstage nach Statusgruppen</b> .....	667
27.3.6	<b>Abwesenheitstage nach Behördengruppen</b> .....	668
27.3.7	<b>Abwesenheitstage nach Geschlecht</b> .....	668
27.3.8	<b>Abwesenheitstage nach Alter</b> .....	668
27.3.9	<b>Gegenüberstellung mit den Abwesenheitszeiten der AOK-Statistik</b> .....	670
	<b>Literatur</b> .....	672
28	<b>Karenztag und Absenkung der Lohnersatzrate – eine ökonomische Einordnung</b> .....	675
	<i>Nicolas R. Ziebarth</i>	
28.1	<b>Einleitung</b> .....	676
28.2	<b>Definition von Karenztagen und Lohnersatzraten</b> .....	678
28.3	<b>Mögliche Folgen bei Einführung eines Karenztages</b> .....	678

## Inhaltsverzeichnis

28.4	<b>Mögliche Folgen der Absenkung der Lohnersatzrate</b>	680
28.5	<b>Handlungsempfehlungen und Fazit</b>	682
	Literatur	683
	<b>Serviceteil</b>	687
	Die Autorinnen und Autoren	688
	Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (10. Revision, German Modification, Version 2025)	708
	Branchen in der deutschen Wirtschaft basierend auf der Klassifikation der Wirtschaftszweige (Ausgabe 2008/NACE)	718
	Glossar	723
	Stichwortverzeichnis	731

# Einführung in das Thema

## Inhaltsverzeichnis

- Kapitel 1 Künstliche Intelligenz –  
eine Schlüsseltechnologie – 3**  
*Hartmut Hirsch-Kreinsen*
- Kapitel 2 Gesünder, produktiver, automatisierter?  
Ethische Aspekte von KI im Betrieblichen  
Gesundheitsmanagement – 15**  
*Lou Therese Brandner*
- Kapitel 3 Datenschutz- und KI-rechtliche Aspekte  
beim Einsatz Künstlicher Intelligenz  
in Unternehmen – 27**  
*Rolf Schwartmann und Moritz Köhler*
- Kapitel 4 „KI als Medizin“ für das Gesundheitswesen? – 39**  
*Martin Roesler*
- Kapitel 5 Von der Personalplanung  
bis zum Gesundheitsmanagement:  
Nutzungsmöglichkeiten von künstlicher  
Intelligenz im Personalwesen – 51**  
*Alexander Dregger*
- Kapitel 6 Künstliche Intelligenz in der betrieblichen  
Prävention und Gesundheitsförderung – 65**  
*David Matusiewicz*

**Kapitel 7 KI in der Betrieblichen Gesundheitsförderung und Prävention – ausgewählte Perspektiven zum internationalen Forschungsstand – 83**

*Martin Lange und Andrea Schaller*



# Künstliche Intelligenz – eine Schlüsseltechnologie

Hartmut Hirsch-Kreinsen

## Inhaltsverzeichnis

- 1.1 Der KI-Boom – 4**
- 1.2 Zur Dynamik der KI – 5**
  - 1.2.1 KI, ein Sammelbegriff – 5
  - 1.2.2 Weitreichende Versprechungen – 6
- 1.3 Diffusion in Grenzen – 7**
  - 1.3.1 Zurückhaltende Verbreitung – 7
  - 1.3.2 Übersehene Anwendungsprobleme – 8
- 1.4 Moderate Folgen für Arbeit – 10**
- 1.5 Perspektiven – 11**
- Literatur – 12**

## ■■ Zusammenfassung

*Thema des Beitrages ist die dynamische Entwicklung der Künstlichen Intelligenz (KI). Diese Dynamik ist ohne Frage mit den schnellen Fortschritten der KI-Technologien erkläbar. Jedoch darf nicht übersehen werden, dass die rasante KI-Entwicklung zugleich von weitreichenden Versprechungen und einer großen Faszination über ihre besondere Leistungsfähigkeit angetrieben wird. Denn versprochen wird nicht nur, dass mit KI große Produktivitätsfortschritte und hohe ökonomische Wachstumsraten erzielt werden, sondern auch, dass vielfältige und drängende gesellschaftliche und ökologische Herausforderungen bewältigt werden können. Daran knüpft der folgende Beitrag an: Erstens wird der Zusammenhang zwischen der dynamischen Technologieentwicklung und den damit einhergehenden weitreichenden Versprechungen diskutiert. Zweitens wird gezeigt, dass trotz weitreichender Versprechungen bislang nur eine begrenzte Diffusion der KI in zentralen gesellschaftlichen Bereichen wie der Industrie und dem Gesundheitswesen beobachtbar ist. Eine Ursache hierfür sind oftmals übersehene technische, ökonomische und soziale Anwendungsprobleme sowie generelle Leistungsgrenzen dieser Technologie. Drittens wird gezeigt, dass in Hinblick auf die Folgen der KI für Arbeit und Beschäftigung derzeit bestenfalls moderate Folgen erkennbar sind und langfristige Konsequenzen kaum valide prognostiziert werden können. Abschließend wird begründet, dass entgegen allen Zukunftsvisionen einer sogenannten menschenähnlichen KI eine anwendungsorientierte Weiterentwicklung der KI eine realistische Perspektive ist.*

### 1.1 Der KI-Boom

Künstliche Intelligenz (KI) bzw. im internationalen Sprachgebrauch Artificial Intelligence (AI) ist seit Jahren Gegenstand eines Booms, ja eines Hypes nicht nur in fachlichen Diskussionen, sondern auch in Politik und Gesellschaft insgesamt. Aufgrund der ihr zugeschriebenen

Fähigkeit, komplexe Probleme „autonom“ und „smart“ zu bearbeiten, löst sie eine große Faszination in der Fachwelt und in der Öffentlichkeit aus. Spätestens seit Ende 2022 das sogenannte Große Sprachsystem „ChatGPT“ einer breiten Öffentlichkeit vorgestellt worden ist, hat der KI-Diskurs nochmals einen Schub erhalten.

Allerdings hat die öffentliche Faszination von KI widersprüchliche Seiten: Einerseits gilt KI aufgrund ihrer besonderen Leistungsfähigkeit als die für die Zukunft unverzichtbare „Schlüsseltechnologie“, um vielfältige und drängende ökonomische und gesellschaftliche Herausforderungen zu bewältigen und um besonders aus deutscher Sicht im globalen Technologiewettlauf nicht ins Hintertreffen zu geraten. Dies ist das zentrale Argument der KI-Strategie der Bundesregierung (Bundesregierung 2018); ähnlich äußerte sich auch schon in der zweiten Hälfte der 1970er Jahre der damalige Forschungsminister Hans Matthöfer (Ahrweiler 1995, S. 85). Andererseits aber begründet diese Faszination auch vielfältige gesellschaftspolitische Befürchtungen über unerwünschte soziale, ethische und ökologische Folgen dieser Technologie. Diese formulieren selbst renommierte KI-Experten wie der Nobelpreisträger und „Godfather“ der KI Geoffrey Hinton (NYT 2023).

Freilich lassen sich Faszination und „Hype“ nicht nur mit der schnellen technologischen Entwicklung der KI begründen. Vielmehr spielen dabei auch die damit verknüpften weitreichenden Versprechungen und Erwartungen über die besondere Leistungsfähigkeit der KI sowie die mit ihrer Hilfe erwarteten ökonomischen Gewinne und die erhoffte Bewältigung gesellschaftlicher und ökologischer Herausforderungen eine entscheidende Rolle. Denn wie die sozialwissenschaftliche Technik- und Innovationsforschung seit langem zeigt, sind Erfolg und Verbreitung neuer Technologien stets verknüpft mit Versprechungen, die bei interessierten Akteuren wie potenziellen Nutzenden, politischen Akteuren, Unternehmen oder auch generell der Öffentlichkeit Erwartungen und

Interessen wecken und vor allem Investitionen in ihre Entwicklung anstoßen. In der Forschung wird dieser Zusammenhang konzeptionell mit dem Begriff der „promising technology“ gefasst (z. B. van Lente und Rip 1998; Borup et al. 2006).

Dies zeigt sich besonders ausgeprägt bei der Entwicklung der KI. Vor allem waren die mit dieser Technologie verknüpften Versprechungen in der Vergangenheit stets die Voraussetzungen dafür, dass Entwicklungsprobleme und Durststrecken in den 1970er und 1990er Jahren, die auch als „AI-Winter“ (Teich 2020) bezeichnet werden, überwunden wurden. Erwartungen etwa bei potenziellen Anwendern und in der Öffentlichkeit wurden stets erneut geweckt und die KI-Dynamik nahm immer wieder Fahrt auf (Hirsch-Kreinsen 2023). Auf den Punkt bringt diese von Versprechungen und Erwartungen geprägte KI-Dynamik ein renommierter Protagonist der amerikanischen KI-Entwicklung und Gründungsmitglied der American Association for Artificial Intelligence, Bruce G. Buchanan: „*The history of AI is a history of fantasies, possibilities, demonstrations, and promise.*“ (Buchanan 2006, S. 53).

Daran knüpft die folgende Argumentation an: Erstens wird die Dynamik der KI in Hinblick auf den engen Zusammenhang zwischen der Technologieentwicklung und den mit ihr einhergehenden weitreichenden Versprechungen diskutiert. Zweitens wird gezeigt, dass gleichwohl bislang nur eine begrenzte Diffusion der KI in zentralen gesellschaftlichen Bereichen wie der Industrie und dem Gesundheitswesen beobachtbar ist. Eine im Diskurs oftmals übersehene Ursache hierfür sind Anwendungsprobleme und Leistungsgrenzen dieser Technologie. Drittens wird gezeigt, dass derzeit bestenfalls moderate Folgen der KI für Beschäftigung und Arbeit erkennbar sind und langfristige Effekte für den Arbeitsmarkt valide nicht prognostiziert werden können. Abschließend wird begründet, dass entgegen allen Visionen einer menschenähnlichen KI, die in Zukunft menschliche Fähigkeiten erreichen oder gar überschreiten wird, eine anwendungs-

orientierte Weiterentwicklung der KI eine realistische Perspektive ist.

Die methodische Basis der Argumentation umfasst eine Zusammenfassung der Ergebnisse empirischer Erhebungen und Recherchen über die Entwicklungsdynamik der KI speziell in der Bundesrepublik, die in den letzten Jahren an der TU Dortmund und der Sozialforschungsstelle Dortmund durchgeführt worden sind (vgl. hierzu Hirsch-Kreinsen 2023; Hirsch-Kreinsen und Krokowski 2023).

## 1.2 Zur Dynamik der KI

### 1.2.1 KI, ein Sammelbegriff

Es ist bis heute schwierig, eine allgemein anerkannte Definition von KI zu finden. Erstmals geprägt wurde dieser Begriff bei einer als „Geburtsstunde der KI“ bezeichneten Konferenz am Dartmouth College, New Hampshire in den USA im Jahr 1956. Einer der Initiatoren dieser Konferenz, der Mathematiker und Informatiker John McCarthy, gilt als der „Gründungsvater“ der KI, der damals den Begriff *Artificial Intelligence* eingeführt hat (Konrad 1998; Koehler 2021).

Zu einer Begriffsklärung kann in Anlehnung an ein renommiertes KI-Lehrbuch festgehalten werden, dass Künstliche Intelligenz eine wissenschaftliche Disziplin ist, „... die das Ziel verfolgt, menschliche Wahrnehmungs- und Verstandesleistungen zu operationalisieren und durch Artefakte, kunstvoll gestaltete technische – insbesondere informationsverarbeitende – Systeme verfügbar zu machen“ (Görz et al. 2021, S. 2). Anders formuliert: Die entwicklungsleitende Idee der KI ist, dass kognitive Leistungen des Menschen als Informationsverarbeitungsprozesse begriffen werden und menschliches Denken als ein formalisierbarer Eingabe-Ausgabe-Prozess anzusehen ist. KI zielt entsprechend darauf, ein solches System mit Hilfe von Computertechnologien technologisch zu reproduzieren und letztlich als technisches System in verschiedensten sozialen und ökonomischen Bereichen anwendbar

zu machen (Ahrweiler 1995). Daher wird die Fähigkeit eines Systems, bestimmte Funktionen und Ziele „intelligent“ und „autonom“ zu erreichen, als entscheidendes Kriterium für eine Definition von KI angesehen.

Genau genommen hat KI aber den Charakter eines Sammelbegriffs (Bauberger et al. 2021). Er umfasst ein weites Feld verschiedener Konzepte und Methoden, die in historisch unterschiedlichen Entwicklungsphasen dieser Technologie ebenso unterschiedliche Relevanz besessen haben. Zu nennen sind hier – sehr verkürzt – Methoden der sog. symbolischen bzw. wissensbasierten KI und darauf basierende anwendungsorientierte Expertensysteme (1970er- und 1980er-Jahre), klassisches Maschinelles Lernen und beginnende Robotik (1990er-Jahre), künstliche Neuronale Netze und Deep Learning (seit den 2010er-Jahren) und aktuell die sog. generativen KI-Systeme bzw. Großen Sprachmodelle. Spätestens mit dem Beginn der zweiten Dekade der 2000er wird eine beschleunigte Dynamik der KI erkennbar, die sich seitdem kontinuierlich verstärkt. Treibende Faktoren hierfür sind die Entwicklung hochkomplexer KI-Methoden wie etwa der Statistik, die über das Internet bereitgestellte massive Zunahme global verfügbarer Daten und die Möglichkeiten, diese mit Big-Data-Methoden zu nutzen, sowie eine exponentiell steigende und sich verbilligende Rechnerleistung. In der Literatur wird dieser technologische Schub nach 2010 treffend als „Big Bang of Deep Learning“ charakterisiert (Görz et al. 2021, S. 9).

Die Methoden des Maschinellen Lernens, insbesondere sog. Deep-Learning-Netze, sind auch die Basis für die seit 2022 intensiv diskutierte generative KI, deren prominentestes Beispiel das System ChatGPT ist. Diese KI ist in der Lage, auf der Basis sehr großer Datenmengen und initiiert durch gezielte Anweisungen oder Fragen, die sog. Prompts, neue Inhalte wie Texte, Bilder, Musik oder Videos zu generieren. Die oft in diesem Zusammenhang erwähnten Großen Sprachmodelle (Large Language Models) können als ein Sonderfall generativer KI angesehen werden, der auf

Textverarbeitung spezialisiert ist. Ein Überblick über die seit 2022 zunehmend große Zahl von Sprachmodellen und deren Leistungsfähigkeit findet sich auf der Plattform Chatbot Arena.<sup>1</sup>

## 1.2.2 Weitreichende Versprechungen

Mit dieser technologischen Entwicklung unmittelbar verknüpft sind weitreichende Versprechungen und vielfältige Erwartungen. Voraussetzung hierfür ist die angesprochene Mehrdeutigkeit des Sammelbegriffs KI. Insbesondere spielt hierbei die gebräuchliche und leicht zugängliche, sich an menschlichen Eigenschaften und Verhaltensweisen orientierende KI-Metaphorik mit Kategorien wie „Information“, „lernend“, „smart“ oder „autonom“ eine einflussreiche Rolle. Denn sie bietet viele scheinbar leicht verständliche Interpretations- und Anschlussmöglichkeiten für ein Verständnis der komplexen Technologie und Wunschvorstellungen über ihre zukünftige Entwicklung.

Der Kern der KI-Versprechungen ist, dass mit KI weitreichende ökonomische Wachstumsperspektiven eröffnet werden. Folgt man einer frühen Prognose aus dem Jahr 2018, so sollte die Nutzung von KI innerhalb von fünf Jahren – also bis Mitte der 2020er Jahre – eine zusätzliche Bruttowertschöpfung in Höhe von ca. 31,8 Mrd. € im produzierenden Gewerbe allein in Deutschland erbringen (Institut für Innovation und Technik 2018). Prognosen zu den ökonomischen Effekten der generativen KI gehen noch deutlich weiter: Einer Studie der Unternehmensberatung McKinsey (Chui et al. 2024) zufolge könnte generative KI pro Jahr einen Wert von 2,6 bis 4,4 Billionen US-Dollar zum globalen BIP beitragen. Dabei wird der generativen KI eine sich schnell beschleunigende Diffusion prognostiziert und es wird erwartet, dass KI im Zusammenspiel mit anderen Technologien massiv die Produktivität

<sup>1</sup> ► <https://lmarena.ai/>.

von Arbeit steigt; gesprochen wird hier von zusätzlichen Steigerungsraten der Produktivität von 0,5 bis 3,4 % jährlich.

Allerdings gehen die Versprechungen deutlich über rein ökonomische Aspekte hinaus und sie haben explizit gesellschaftsbezogenen Charakter. So wird in der Strategie Künstliche Intelligenz der Bundesregierung (Bundesregierung 2018) ein wünschenswerter Strukturwandel von Arbeit und Arbeitsmarkt als Ziel der KI-Einführung angesprochen. KI biete die Voraussetzung, Bildung sowie Aus- und Weiterbildung zu unterstützen und damit auch den Fachkräftemangel zu reduzieren. Hervorgehoben wird zudem, KI werde geradezu zwangsläufig zu guter Arbeit, anspruchsvollen Jobs und zu neuen Arbeitsformen führen. Insgesamt biete KI die Voraussetzung, Arbeit anspruchsvoller zu gestalten und massiv aufzuwerten und neue und qualifizierte Jobs zu schaffen (Bitkom und DFKI 2017).

Darüber hinaus soll mithilfe von KI die Bewältigung der vielfältigen ökologischen Herausforderungen erleichtert werden. So wird in der deutschen KI-Strategie betont, dass KI erhebliche Chancen für den Schutz von Umwelt, Klima, Ressourcen und den Erhalt der Artenvielfalt mit sich bringe (z. B. BMU 2021; Peteranderl 2021). Auch wird das Versprechen formuliert, mittels KI könne der Energie- und Ressourcenverbrauch optimiert, ja reduziert werden (z. B. Enquetekommission 2020). Sehr zugespitzt bekräftigt dieses weitreichende Versprechen letztens Sam Altman, prominenter Entwickler von ChatGPT und Vorsitzender des globalen KI-Unternehmens Open-AI, wonach mit KI die ökologische Krise der Menschheit gelöst werden könnte (FAZ 2025).

### 1.3 Diffusion in Grenzen

Angesichts der KI-Dynamik und der weitreichenden Versprechungen liegt die Vermutung nahe, dass diese eine schnelle Diffusion und breite Anwendung von KI-Systemen nach sich ziehen. Indes zeigen die vorliegenden empirischen Daten sehr divergierende und wenig

eindeutige Tendenzen der KI-Diffusion. Mit einer hohen Dynamik diffundieren KI-Systeme in den letzten Jahren weltweit vor allem in verschiedenste Dienstleistungssektoren, in IT-, Kommunikations- und Informationsbereiche sowie in viele alltäglich genutzte Systeme. So werden neuronale Netze von Onlineshops für Kaufempfehlungen und von sozialen Netzwerken für Gesichtserkennung oder das Sortieren und Optimieren von Fotos genutzt. Weitere seit Jahren bekannte Anwendungen sind Spamfilter, Suchmaschinen, maschinelle Übersetzungen und Systeme der Spracherkennung. KI-Anwendungen finden sich auch „implizit“ als Komponente innerhalb von Systemen der Sprach- und Mustererkennung, sodass sie kaum als solche direkt identifizierbar sind. Vor allem sind die „Social Media“ ohne die Nutzung von KI zur Muster- und Bilderkennung oder für personalisierte Werbung schon seit Jahren nicht mehr denkbar. Insbesondere aber ist auf eine geradezu atemberaubende Verbreitung generativer Systeme zu verweisen, die etwa im Wissenschaftsbereich, im Journalismus, in der Werbebranche oder im alltäglichen Gebrauch als inzwischen breit genutzte komfortable Suchmaschinen intensiv genutzt werden (Zysman und Nitzberg 2024).

#### 1.3.1 Zurückhaltende Verbreitung

Demgegenüber ist in vielen Wirtschafts- und Gesellschaftssektoren, die weit heterogener und differenzierter strukturiert sind als etwa IT-, Kommunikations- und Informationssektoren, eine bislang sehr zögerliche Anwendung von KI zu beobachten. Ganz offensichtlich fallen in vielen Fällen die reale Ebene einer KI-Nutzung und die Diskursebene mit ihren Versprechen weit auseinander. Dies gilt insbesondere für so zentrale gesellschaftliche Bereiche wie den industriellen Sektor und das Gesundheitswesen.

Der industrielle Sektor Deutschlands steht unter nachhaltigen Rationalisierungs- und Modernisierungsdruck und es ist seit langem unstrittig, dass eine möglichst breite Nutzung di-

gitaler Technologien hierfür ein unverzichtbares Mittel ist (z. B. Forschungsumunion und acatech 2013). Indes zeigen neuere Studien nicht nur einen zurückhaltenden Einsatz von digitalen Technologien generell, sondern besonders auch von KI-Komponenten und Systemen. Einer Studie des Fraunhofer Instituts für System- und Innovationsforschung zufolge finden sich im Jahr 2022 fortgeschrittene digitale Systeme lediglich bei Großunternehmen und Ausrüsterunternehmen aus der Investitionsgüterindustrie, während bei den meisten mittleren und kleineren Betrieben von einer Stagnation und teilweise einem Rückgang des Digitalisierungsniveaus gesprochen werden kann (Lerch et al. 2024). Dabei zeigen sich allerdings deutliche Unterschiede zwischen unterschiedlichen Betriebsgrößen. Während vor allem größere Betriebe mit über 500 Beschäftigten mit KI experimentieren, finden sich solche Systeme nur in Ausnahmefällen bei mittleren und kleineren Betrieben. Als zukünftig potenzielle industrielle Anwendungsfelder von KI gelten beispielsweise Funktionen der Qualitätskontrolle, der Produktionssteuerung und Fabrikplanung, der Logistik, der Instandhaltung und auch des Human Resource Managements (Heimberger et al. 2024). Andere Untersuchungen bestätigen diese Befunde. So heißt es, dass sich die Einführung von KI im verarbeitenden Gewerbe vielerorts noch in den „Startlöchern“ befindet und bestenfalls KI-Anwendungen im Experimentier- und Pilotstadium anzutreffen seien (Forschungsbeirat und acatech 2024).

Ein ähnliches Bild zeigt sich im Gesundheits- und Medizinsektor. Zwar wird für diesen Sektor KI als „Allheilmittel“ für eine Vielzahl gegenwärtiger und zukünftiger Probleme angesehen. Es wird erwartet, dass digitale Lösungen das Gesundheitssystem effizienter, weniger bürokratisch und zuverlässiger machen sowie seine Qualität steigern (Graßmann und Pfeiffer 2024, S. 472). Im Zentrum stehen hier Anwendungsbereiche wie: erstens die digitale Dokumentation, Robotik- und Assistenzsysteme oder digitale bzw. virtuelle Kommunikation. Als ein wei-

terer Schwerpunkt des Einsatzes von KI und digitaler Systeme generell wird seit längerem der Pflegebereich angesehen (z. B. Melzer et al. 2022). Insgesamt gesehen steht aber neuesten Studien zufolge die Verbreitung von KI in diesen und weiteren medizinischen Funktionsbereichen – wenn überhaupt – erst am Anfang und es kann bestenfalls von einer selektiven KI-Nutzung gesprochen werden (Gerber 2025).

### 1.3.2 Übersehene Anwendungsprobleme

Frage man nach den Gründen für diese insgesamt widersprüchliche Situation zwischen einerseits hochfliegenden Versprechungen und andererseits einer nur zögerlichen Nutzung der KI in zentralen gesellschaftlichen und wirtschaftlichen Bereichen, so ist auf Anwendungsprobleme und Leistungsgrenzen der KI zu verweisen, die im Diskurs oft übersehen werden. Sehr vereinfacht sind hier die folgenden hervorzuheben:

#### ■ Schwieriger Wissenstransfer und Fachkräftemangel

Als zentrales Problem einer erfolgreichen KI-Einführung ist der Wissenstransfer zwischen KI-Experten einerseits und den Fachleuten aus den potenziellen Anwendungsbereichen andererseits. Denn insbesondere KI-Systeme und die gegebenen technisch-organisatorischen Bedingungen eines Anwendungsbereichs bedürfen stets der wechselseitigen Abstimmung, die allein durch die enge Kooperation beider Seiten gelingen kann. Indes trifft dies oftmals auf organisatorische Beharrungskräfte und eine wenig digital ausgerichtete Mentalität bei potenziellen Anwenderorganisationen, sodass der erforderliche Wissenstransfer kaum gelingt. Beispielsweise wird von Fachleuten immer das vorherrschende Denken in „Silos“ beklagt. Mitarbeitende seien oftmals nicht bereit, ihre Arbeitsbereiche für abteilungsübergreifende Kommunikation und Wissenstransfer zu öffnen. Hinzu kommt das

seit langem thematisierte Problem fehlender Informatik-, Daten- und KI-Experten wie etwa Datenanalysten, Data Scientists oder auch sog. Data Engineers (Ecker et al. 2021; Hirsch-Kreinsen 2023).

### ■■ Ökonomische Risiken

In der Regel verbinden sich mit der Einführung von KI-Systemen entgegen aller Gewinnerwartungen nur schwer kalkulierbare ökonomische Risiken. So weisen Praktiker vor allem auf die hohen Investitions- und Implementationskosten der neuen Systeme und ihre unklaren Rentabilitätsaussichten hin (Forschungsbeirat und acatech 2022). Zudem wird vielfach auf den häufig unklaren Aufwand und Folgeinvestitionen hingewiesen, die notwendig seien, um die Systeme wirklich zum Laufen zu bringen. Es werde oft unterschätzt, so Fachleute, dass ein KI-System „nur einen Bruchteil einer KI-Lösung ausmacht.“ Hingegen könne der Aufwand für die Erfassung und Aufbereitung der notwendigen Daten bis zu 80 % der Kosten umfassen. Dabei geht es vor allem darum, die verfügbaren Daten zu analysieren, nutzbar zu machen und die für einen effektiven Systembetrieb erforderliche Datenqualität dauerhaft zu sichern (Keim und Sattler 2020; Ecker et al. 2021).

### ■■ Datenprobleme

Ungeklärt bleibt in vielen Fällen darüber hinaus die Frage, ob die Anwendenden überhaupt über hinreichend valide Daten für eine effektive Nutzung von KI-Systemen verfügen. Zwar kann man in der Regel von einer großen Zahl verfügbarer, aber bislang ungenutzter Daten in Unternehmen oder auch im Gesundheitswesen ausgehen. Jedoch stellt sich oftmals erst im konkreten Anwendungsfall heraus, ob Qualität und Validität der verfügbaren Daten für die angestrebten Systemfunktionen überhaupt ausreichend sind (Hirsch-Kreinsen 2023, S. 158 ff.). Oft handelt es sich dabei um das Problem, dass nur geringe Datenmengen – „little data“ im Unterschied zu „big data“ – für einen effektiven Einsatz von KI-Systemen verfügbar sind. Dies ist in der Regel in in-

dustriellen Prozessen der Fall, wo nur geringe Mengen von qualitativ schwer einschätzbareren Daten zur Verfügung stehen. Die Folge ist beispielsweise, dass eine von einem System erwünschte Prognosewahrscheinlichkeit nur mit einem schnell steigenden Daten- und Kostenaufwand erreicht werden kann, der sowohl das technisch Machbare als auch das ökonomisch Sinnvolle überschreitet (Pfeiffer 2023, S. 389).

Letztlich lässt sich aber auch nicht jenes Anwendungsrisiko vermeiden, das im kritischen Diskurs über die Sprachsysteme als „Halluzinationen“ bezeichnet wird. Gemeint sind damit irreführende oder falsche Aussagen der Systeme. Ihre Ursachen sind in der Fachdebatte umstritten, vielfach aber werden sie als schwer vermeidbares statistisches Problem angesehen und auch auf unzureichende Trainingsdaten zurückgeführt (z. B. Chomsky et al. 2023; Ji et al. 2023). Im Medizinbereich wurde jüngst eine aus solchen Gründen nur eingeschränkte und fehlerhafte Funktionsweise von KI-basierten Diagnosesystemen gezeigt, die auf der Basis begrenzter Datensätze zu falschen Diagnosen kommen (z. B. Hautz et al. 2025). Ähnliches zeigen Befunde einer neuen Untersuchung KI-gestützter Verfahren der medizinischen Bilderkennung. Danach fokussiert sich die KI oft auf irrelevante Bilddetails und kommt so zu falschen Diagnosen und Vorhersagen. Als Grund hierfür wird das sogenannte Unüberwachte Lernen identifiziert, bei dem KI-Systeme auf der Basis sehr großer Datensätze trainiert werden, aus denen die Systeme autonom Muster und Zusammenhänge identifizieren sollen, was aber zu fehlerhaften Ergebnissen führt (Kauffmann et al. 2025).

Fehlentscheidungen und Fehlfunktionen von KI-Systemen können allerdings auch deshalb nicht verhindert werden, weil die Funktionsweise von KI-Modellen nicht zuverlässig ist und kaum erklärbar ist. Ursache hierfür ist das schon seit den 1990ern diskutierte Problem der sogenannten *Explainable Artificial Intelligence*. Denn es gibt für die Funktionsweise Neuronaler Netze keine theoretisch fundierten Erkenntnisse. Aber je unverzichtbarer deren Entscheidungen für menschliches Handeln

werden, desto wichtiger wird es zu verstehen, was die Systeme eigentlich tun. Ein instruktives Beispiel hierfür ist das oben erwähnte KI-unterstützte Diagnosesystem, dessen automatisch erstellte Diagnosen für Patienten weitreichende Konsequenzen haben können, wenn ihnen unkritisch gefolgt wird. In der Fachwelt ist es bis heute umstritten, ob das Problem der Explainability in Zukunft lösbar sein wird (z. B. Brödner 2019; Deepak 2023). Auf diese Probleme und damit verbundene Risiken der KI-Nutzung verweisen nachdrücklich renommierte internationale KI-Experten in einem Anfang 2025 publizierten „International AI Safety Report“ und warnen vor den davon ausgehenden Risiken und Gefahren einer nicht beherrschbaren und unregulierten Anwendung von fortgeschrittenen KI-Systemen (Bengio et al. 2025).

#### 1.4 Moderate Folgen für Arbeit

Trotz der bislang nur begrenzten Diffusion der KI ist ein intensiv diskutiertes Thema, welche Konsequenzen diese Technologie für die Arbeitswelt und ihre Beschäftigten hat. Denn unüberhörbar sind Prognosen wie vor allem Befürchtungen über einen zukünftig weitreichenden Wandel von Arbeit durch KI. So prognostiziert der Internationale Währungsfonds, dass absehbar nicht nur weltweit 40 % aller Arbeitsplätze mit KI befasst sein werden, sondern dass ein wesentlicher Teil vollständig substituiert werden wird. Hingegen werden sich die Anforderungen an Kompetenzen und Tätigkeiten der verbleibenden Arbeitsplätze dramatisch verschieben (Cazzaniga et al. 2024). Ähnliches prognostiziert eine Studie der Beratungsgesellschaft McKinsey, wonach zwischen 2030 und 2060 etwa die Hälfte aller heutigen Arbeitsaktivitäten automatisiert werden könnten. Dies betreffe insbesondere Berufe mit höheren Einkommen und Bildungsniveaus (Chui et al. 2024). Mit anderen Worten, es wird erwartet, dass in Zukunft nicht mehr nur manuelle und industrielle Tätigkeiten, sondern auch viele Nicht-Routine-tätigkeiten, die wissensintensiv und kreativ sind, durch KI-

Systeme ersetzt werden (Zysman und Nitzberg 2024).

Freilich sind solche Prognosen und die damit einhergehenden Befürchtungen methodisch mit einem hohen Unsicherheitsgrad behaftet. Denn sie basieren lediglich auf einer Einschätzung der möglichen Substitutionspotenziale der KI und weniger auf realen Trends. Denn ob ein Arbeitsplatz tatsächlich technologisch substituiert wird, ist vielmehr abhängig von betrieblichen Entscheidungsprozessen und einer Vielzahl von diese beeinflussten ökonomischen und politischen Faktoren, die nur schwer zu prognostizieren sind. Insbesondere darf auch nicht übersehen werden, dass aufgrund der Anwendungsprobleme und Leistungsgrenzen von KI-Systemen vielfältige Barrieren für den Nutzung von KI existieren.

Angesichts dessen spricht vieles dafür, dass derzeit mit KI nur in Ausnahmefällen ein grundlegender Wandel von Arbeit verbunden sein wird. Ähnlich wie eine ganze Reihe von Studien über die Arbeitsfolgen digitaler Technologien generell belegen, ist auch im Fall von KI zumeist nur von einem moderaten Wandel von Arbeit auszugehen. Denn es findet zumeist eine nur schrittweise Adoption der neuen Technologien in Betrieben, im Gesundheitswesen oder auch in Verwaltungen innerhalb der gegebenen technologischen, organisationsstrukturellen und personellen Bedingungen statt. Der Hauptgrund hierfür ist, dass dadurch Risiken und unerwartete Kosten minimiert sowie schrittweise Erfahrungen mit KI gesammelt werden können (z. B. Hirsch-Kreinsen 2020; Heinlein und Huchler 2024). Einerseits impliziert dies, dass bisherige Automatisierungstrends fortgeschrieben werden und einfache und routinähafte Arbeitsprozesse wie etwa in der Logistik, der Montage oder bei einfachen Büroarbeiten fortschreitend durch KI-basierte Steuerungssysteme automatisiert und substituiert werden. Andererseits können bei vielen, etwa qualifizierten Tätigkeiten KI-Systeme als Assistenzsysteme genutzt werden und dadurch das Aufgabenfeld erweitern und vor allem auch Lernprozesse unterstützen. Dies trifft im Industriebereich etwa

auf Tätigkeiten der Instandhaltung oder Überwachung komplexer Anlagen zu. Im medizinischen Bereich kann dies durch die erwähnten KI-gestützten Diagnosesysteme oder die KI-Assistenz bei Operationen geschehen.

Indes ist in der sozialwissenschaftlichen Arbeitsforschung umstritten, dass gerade beim Einsatz von avancierten KI-Systemen menschliche Arbeit auch in Zukunft ein unverzichtbares Flexibilitätspotenzial in Arbeitsprozessen gleicher Art bleibt (Böhle 2023; Pfeiffer und Schrape 2023). Ein Grund hierfür sind die erwähnten Anwendungsprobleme und Leistungsgrenzen der KI, die insbesondere angesichts steigender Anforderungen an Arbeitsprozesse oft nur eine begrenzte Digitalisierung von Arbeitsprozessen ermöglichen und daher „analoge“ Arbeit notwendig bleibt. Darüber hinaus stellen kommunikative und interaktive Tätigkeiten etwa im Servicebereich oder auch im Gesundheitswesen aufgrund ihres situativ-flexiblen und kaum im Voraus kalkulierbaren Charakters auch in Zukunft „basic bottlenecks“ für ihre weitreichende Digitalisierung durch KI dar (Frey und Osborne 2024).

## 1.5 Perspektiven

Fragt man nach den Zukunftsperspektiven der KI-Entwicklung, so sollte man zunächst einmal auf die weitreichenden KI-Versprechungen zurückzukommen. Als technologische Voraussetzung für deren zukünftige Realisation und eine Überwindung der Leistungsgrenzen werden häufig „Durchbrüche“ bei der KI-Entwicklung angesehen (z. B. Armbruster 2022). Denn, so das Versprechen, Systeme wie ChatGPT seien in der Lage „zu lernen“ und sie werden von ihren Entwicklern kontinuierlich aufwendig „nachjustiert“. Formuliert wird damit die Erwartung, über kurz oder lang gebe es eine Maschinenintelligenz, die menschliche Intelligenz übertrifft und die in der Lage ist, sich autonom und ohne menschliches Dazutun ständig weiterzuentwickeln (vgl. insbesondere Kurzweil 2024).

Freilich übersieht diese Perspektive die wachsenden Probleme, ausreichend Daten verfügbar zu haben, die rasant steigenden Kosten einer fortlaufenden Skalierung der bestehenden Systeme sowie den damit einhergehenden massiven Energiebedarf (Bengio et al. 2025). Vor allem aber ist zu bezweifeln, dass eine maschinelle Intelligenz grundsätzlich jemals menschlicher Intelligenz überlegen sein kann. Denn es handelt sich bei KI gleich welcher Konzeption um wenig mehr als um statistische Systeme. Ihre vermeintliche Intelligenz hat den Charakter einer Prognose, basierend auf Daten der Vergangenheit und mit einer in der Praxis unbekannten Prognosewahrscheinlichkeit (Pfeiffer 2023, S. 37). Das bedeutet, dass die Systeme keinerlei „Bewusstsein“, „Vernunft“, kein Bezug zu „Erfahrung und Alltagswissen“ oder gar ein moralisches Verständnis von „richtig“ und „falsch“ haben können. Auch sind die Systeme nicht in der Lage, Abstraktionen zu leisten und Kausalitäten nachzuvollziehen und sich dem zu nähern, was Menschen in besonderer Weise auszeichnet, nämlich Handlungen zu verstehen und zu erklären (Görz et al. 2021, S. 10).

Realistischer ist daher eine Perspektive, die als „Veralltäglichung“ der KI gefasst werden kann. Damit wird eine Entwicklung bezeichnet, durch die die Anwendungsgrenzen der KI evolutionär hinausgeschoben und durch immer leistungsfähigere, für verschiedene Anwendungen speziell ausgelegte Systeme überwunden werden. Einmal betrifft dies die Systeme der generativen KI, die auf der Basis steigender Investitionen in Rechnerleistungen, Speicherkapazitäten und Datennutzung weiter in die Breite diffundieren und für eine schnell wachsende Zahl alltäglicher Anwendungen genutzt werden können. Zum anderen handelt es sich dabei aber auch um Systemkonzepte, die im kleineren Maßstab unterschiedlichste KI-Methoden in sog. hybriden KI-Systemen sehr anwendungsspezifisch nutzen. Als instruktives Beispiel hierfür kann das Anfang 2025 in China vorgestellte generative System DeepSeek angesehen werden. Offensichtlich werden mit diesem System nicht nur sehr ver-

schiedene KI-Methoden kombiniert, sondern es erlaubt auch in besonderer Weise eine für verschiedenste Bereiche anwendungsorientierte Auslegung (Krüger 2025). Eine Folge einer solchen Weiterentwicklung der KI wird daher sein, dass ihre bislang nur zögerliche Anwendung in vielen gesellschaftlichen Sektoren wie in der Industrie und der Medizin beschleunigt und die Leistungsfähigkeit der KI gezielt und an die jeweiligen Anwendungsbedingungen angepasst genutzt werden kann.

Kaum zureichend beantworten lässt sich indes die Frage, welche gesellschaftlichen Folgen mit dieser Perspektive der KI langfristig verbunden sein werden. In jedem Fall aber sind weder weitreichende positive Versprechungen realistisch noch dystopische Negativ-erwartungen etwa in Hinblick auf erhebliche Arbeitsplatzverluste sonderlich überzeugend. Statt Schreckgespenster an die Wand zu malen, sollte über die erwähnten realen Defizite und Anwendungsrisiken der KI gesprochen werden und es sollte darum gehen, diese absehbar zu regulieren (Bengio et al. 2025). Dabei ist vor allem zu sehen, dass die Nutzung von KI und der damit einhergehende soziale und gesellschaftliche Wandel keinesfalls einen „technikdeterministisch“ zwangsläufigen Charakter haben. Vielmehr sind diese stets bestimmt von unternehmerischen, politischen und normativen, generell sozialen Entscheidungen darüber, wie diese Technologie eingesetzt und genutzt werden soll.

## Literatur

- Ahrweiler P (1995) Künstliche Intelligenz-Forschung in Deutschland. Die Etablierung eines Hochtechnologie-Fachs. Waxmann, Münster New York
- Arnbruster A (2022) Große Künstliche Intelligenz. Interview mit Jonas Andrulis FAZ, 17.05.2022. <https://zeitung.faz.net/faz/unternehmen/2022-05-16/75949fd8e1ead69fb32e0132dd78f8c/?GEPC=s5>. Zugegriffen: 17. Mai 2022
- Bauberger S, Beck B, Burchardt A, Remmers P (2021) Ethische Fragen der Künstlichen Intelligenz. In: Görz G et al (Hrsg) Handbuch der Künstlichen Intelligenz, 6. Aufl. De Gruyter, Berlin Boston, S 905–934
- Bengio Y et al (2025) International AI Safety Report. DSIT 2025/001. <https://www.gov.uk/government/publications/international-ai-safety-report-2025>. Zugriffen: 17. Febr. 2025
- Bitkom, DFKI (2017) Künstliche Intelligenz: Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung. Positionspapier. <https://www.bitkom.org/Bitkom/Publikationen/Entscheidungsunterstuetzung-mit-KuenstlicherIntelligenz.html>. Zugriffen: 25. Juni 2021
- BMU – Bundesumweltministerium (2021) Unsere Förderinitiative „KI-Leuchttürme für Umwelt, Klima, Natur und Ressourcen“. <https://www.bmu.de/themen/nachhaltigkeit-digitalisierung/digitalisierung/unsere-foerderinitiative-ki-leuchttuerme>. Zugriffen: 5. Sept. 2021
- Böhle F (2023) Subjektivierendes Arbeitshandeln. In: Bohn R, al (Hrsg) Lexikon der Arbeits- und Industriesoziologie. Nomos, Baden-Baden, S 336–339
- Borup M, Brown N, Konrad K, van Lente H (2006) The sociology of expectations in science and technology. Technol Analysis Strateg Manag 18:285–298
- Brödner P (2019) Grenzen und Widersprüche der Entwicklung und Anwendung ‚Autonomer Systeme‘. In: Hirsch-Kreinsen H, Karacic A (Hrsg) Autonome Systeme und Arbeit. Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt. transcript, Bielefeld, S 69–100
- Buchanan BG (2006) A (Very) Brief History of Artificial Intelligence. Ai Mag 26:53–60
- Bundesregierung (2018) Strategie Künstliche Intelligenz der Bundesregierung. <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>. Zugriffen: 27. Jan. 2025
- Deepak P (2023) ChatGPT is not OK! That's not (just) because it lies. Ai Soc. <https://doi.org/10.1007/s00146-023-01660-x>
- Cazzaniga M et al (2024) Gen-AI: Artificial Intelligence and the Future of Work. IMF Staff Discussion Note SDN2024/001. International Monetary Fund, Washington, DC
- Chomsky, Noam, Roberts, Ian & Watumull, Jeffrey (2023) The False Promise of ChatGPT. In: The New York Times, state: 08 March 2023. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>. Zugriffen: 10. März 2023
- Chui M et al (2024) The economic potential of generative AI. The next productivity frontier. McKinsey
- Ecker W, Coulon C-H, Kohler M (2021) KI in die Anwendung bringen – Eine Gemeinschaftsaufgabe für Hochschulen, Forschungseinrichtungen, Unternehmen und Politik. München. Whitepaper aus der Plattform Lernende Systeme
- Enquetekommission (2020) (BT-Drucks. 19/23700) Bericht der Enquete-Kommission Künstliche Intelligenz

## Kapitel 1 · Künstliche Intelligenz – eine Schlüsseltechnologie

- Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale. Berlin
- FAZ – Frankfurter Allgemeine Zeitung (2025) Übt Demut vor der Maschine! <https://www.faz.net/aktuell/feuilleton/debatten/kuenstliche-intelligenz-sam-altman-an-der-tu-berlin-110285518.html>. Zugegriffen: 11. Febr. 2025
- Forschungsbeirat der Plattform Industrie 4.0, acatech
  - Deutsche Akademie der Technikwissenschaften (2022) Künstliche Intelligenz zur Umsetzung von Industrie 4.0 im Mittelstand
- Forschungsbeirat der Plattform Industrie 4.0., acatech
  - Deutsche Akademie der Technikwissenschaften (2024) Künstliche Intelligenz und industrielle Arbeit
  - Perspektiven und Gestaltungsoptionen. Deutsche Akademie der Technikwissenschaften, München
- Forschungsumun, acatech – Deutsche Akademie der Technikwissenschaften (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0. Plattform Industrie 4.0, Frankfurt am Main
- Frey CB, Osborne M (2024) Generative AI and the future of work: a reappraisal. Brown J World Aff 30:1–17
- Gerber C (2025) Digitalisierung im Gesundheitswesen: Status quo, Herausforderungen und Folgen für die Arbeitsbedingungen. Discussion Paper SP III 2025-301. Wissenschaftszentrum Berlin für Sozialforschung, Berlin
- Görz G, Braun T, Schmid U (2021) Einleitung. In: Görz G, Braun T, Schmid U (Hrsg) Handbuch der Künstlichen Intelligenz, 6. Aufl. De Gruyter, Berlin Boston, S 1–26
- Graßmann S, Pfeiffer S (2024) AI and medicine: separating fact from fiction in the quest for effective and high quality healthcare. In: Heinlein M, Huchler N (Hrsg) Artificial intelligence in society. Springer, Wiesbaden, S 465–491
- Hautz WE et al (2025) Diagnoses supported by a computerised diagnostic decision support system versus conventional diagnoses in emergency patients (DDX-BRO): a multicentre, multiple-period, double-blind, cluster-randomised, crossover superiority trial. Lancet Digit Health 7:e136–e144
- Heimberger H, Jäger A, Maloca S (2024) Künstliche Intelligenz in der Produktion. Mitteilungen aus der ISI-Erhebung Modernisierung in der Produktion 83. FhG ISI Karlsruhe
- Heinlein M, Huchler N (Hrsg) (2024) Künstliche Intelligenz, Mensch und Gesellschaft. Springer, Wiesbaden
- Hirsch-Kreinsen H (2020) Digitale Transformation von Arbeit. Entwicklungstrends und Gestaltungsansätze. Kohlhammer, Stuttgart
- Hirsch-Kreinsen H (2023) Das Versprechen der Künstlichen Intelligenz. Gesellschaftliche Dynamik einer Schlüsseltechnologie. Campus, Frankfurt New York
- Hirsch-Kreinsen H, Krokowski T (2023) Technologien versprechen Künstliche Intelligenz. Berlin J Soziol 33:453–485
- Institut für Innovation und Technik (2018) Potenziale der Künstlichen Intelligenz im Produzierenden Gewerbe in Deutschland. Institut für Innovation und Technik in der VDI/VDE Innovation + Technik, Berlin
- Ji Z et al (2023) Survey of hallucination in natural language generation. ACM Comput Surv 55:1–38
- Kauffmann J, Dippel J, Ruff L et al (2025) Explainable AI reveals Clever Hans effects in unsupervised learning models. Nat Mach Intell 7:412–422
- Keim D, Sattler K (2020) Von Daten zu KI. Intelligentes Datenmanagement als Basis für Data Science und den Einsatz lernender Systeme. München. Whitepaper aus der Plattform Lernende Systeme
- Koehler J (2021) Zum Begriff der Künstlichen Intelligenz. Preprint Version. [https://www.dfk1.de/fileadmin/user\\_upload/import/11269\\_KI-Kunst-JKoehler-PreprintVersion.pdf](https://www.dfk1.de/fileadmin/user_upload/import/11269_KI-Kunst-JKoehler-PreprintVersion.pdf). Zugegriffen: 6. März 2025
- Konrad E (1998) Zur Geschichte der Künstlichen Intelligenz in der Bundesrepublik Deutschland. In: Siekes D, et al (Hrsg) Sozialgeschichte der Informatik. Studien zur Wissenschafts- und Technikforschung. Springer, Wiesbaden, S 287–296
- Krüger A (2025) Ist die neue KI der Chinesen ein echter Durchbruch, Herr Krüger? FAZ-net. <https://zeitung.faz.net/faz/unternehmen/2025-02-01/b6860970a9992edeb447ec81b25d1f8b?GEPC=s5>. Zugegriffen: 6. Febr. 2025
- Kurzweil R (2024) The singularity is nearer: when we merge with AI. Penguin, New York
- van Lente H, Rip A (1998) Expectations in technological developments: an example of prospective structures to be filled in by agency. In: Cornelis DC, van der Meulen B (Hrsg) Getting new technologies together: studies in making sociotechnical order. Springer, Berlin New York, S 203–229
- Lerch C, Jäger A, Horvat D (2024) (r)Evolution 4.0: Auf den Spuren des rückläufigen I4.0-Fortschritts. Modernisierung der Produktion. Mitteilungen aus der ISI Erhebung. Ausgabe 82. FhG ISI Karlsruhe
- Melzer M, Rösler U, Schlicht L (2022) Digitale Transformation personenbezogener Arbeit – am Beispiel der professionellen Pflege. In: Bamberg E, Ducki A, Janneck M (Hrsg) Digitale Arbeit gestalten. Springer, Berlin New York, S 147–166
- NYT – New York Times (2023) ‘The Godfather of A.I.’ Leaves Google and Warns of Danger Ahead. <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html>. Zugegriffen: 27. Jan. 2025
- Peteranderl S (2021) Wie künstliche Intelligenz die Klimakrise bekämpfen kann. Der Spiegel. [https://www.spiegel.de/ausland/kuenstliche-intelligenz-wie-tech-werkzeuge-im-kampf-gegen-den-klimakrise-helfen-a-f7ed1f82-790a-4d4f-be34-22196d99730e?sara\\_ecid=soci\\_upd\\_wbMbjhOSvViISjc8RPtU89NcCvtlFcJ](https://www.spiegel.de/ausland/kuenstliche-intelligenz-wie-tech-werkzeuge-im-kampf-gegen-den-klimakrise-helfen-a-f7ed1f82-790a-4d4f-be34-22196d99730e?sara_ecid=soci_upd_wbMbjhOSvViISjc8RPtU89NcCvtlFcJ). Zugegriffen: 17. Aug. 2021

- 1
- Pfeiffer S (2023) KI im Unternehmen – Herausforderung an die betriebliche Gestaltung moderner Arbeit. DGUV Forum 11/2023.
- Pfeiffer S, Schrape J-F (2023) Digitale Transformation der Arbeit. In: Bohn R, al (Hrsg) Lexikon der Arbeits- und Industriesoziologie. Nomos, Baden-Baden, S 134–137
- Teich IB (2020) Meilensteine der Entwicklung Künstlicher Intelligenz. Informatik Spektrum 43:276–284
- Zysman J, Nitzberg M (2024) Generative AI and the future of work: augmentation or automation? Weizenbaum Discussion Paper 38



# Gesünder, produktiver, automatisierter? Ethische Aspekte von KI im Betrieblichen Gesundheitsmanagement

*Lou Therese Brandner*

## Inhaltsverzeichnis

- 2.1      Einleitung: KI zwischen Dystopie und Allheilmittel – 16
- 2.2      KI-Ethik als Forschungsfeld – 16
- 2.3      Fairness und Antidiskriminierung:  
Ist KI objektiv? – 17
- 2.4      Privatheit versus Transparenz: Werte im Konflikt? – 19
- 2.5      Menschliche Autonomie – Aufsicht und  
Letztentscheidung – 21
- 2.6      Fazit – 23
- Literatur – 23

## ■■ Zusammenfassung

*Der Einsatz von KI-Technologien im Betrieblichen Gesundheitsmanagement hat das Potenzial, die Gesundheit und Sicherheit von Beschäftigten z. B. durch innovative Präventions- und Interventionsmaßnahmen zu fördern. Gleichzeitig birgt KI erhebliche Risiken für das individuelle und kollektive Wohlergehen, wie die Diskriminierung marginalisierter Gruppen, Eingriffe in die Privatheit und informationelle Selbstbestimmung sowie den Verlust menschlicher Autonomie und Kontrolle. Dieser Beitrag beleuchtet verschiedene ethische Aspekte von KI – Fairness, Privatheit, Transparenz, menschliche Aufsicht und Letztentscheidung – im Kontext der betrieblichen Gesundheit. Anhand von Beispielen wie der KI-gestützten Diagnostik, dem Tracking von Beschäftigten oder der Analyse von Affekten wird illustriert, dass die Einbeziehung ethischer Gesichtspunkte essenziell ist für die Entwicklung und den Einsatz gesellschaftlich verträglicher KI-Technologie.*

### 2.1 Einleitung: KI zwischen Dystopie und Allheilmittel

Der vermehrte Einsatz von künstlicher Intelligenz (KI) ruft im öffentlichen Diskurs sowohl messianische Erwartungen als auch dystopische Ängste hervor. Während Anbieter von KI-basierten Systemen die Technologie als Lösung für alle möglichen Probleme von Armut (Walther 2024) über Krebsleiden (Cancer Research Institute 2024) bis hin zum Klimawandel (Masterson 2024) anpreisen, stehen auf der anderen Seite Befürchtungen hinsichtlich computergetriebener Diskriminierung, der **Ersetzung menschlicher Arbeitskräfte** oder gar der Entstehung einer übermächtigen Superintelligenz. Ähnliches gilt für den Einsatz von KI im Gesundheitswesen: Manche sehen in KI das Allheilmittel für eine Vielzahl an Herausforderungen, z. B. eine verbesserte Diagnosengenauigkeit oder optimierte Arbeitsabläufe. Die Internationale Arbeitsorganisation (ILO 2023) schätzt, dass weltweit jährlich 395 Mio.

Personen bei der Arbeit verunfallen, während fast drei Millionen Personen aufgrund von arbeitsbedingten Unfällen und Erkrankungen ihr Leben verlieren; KI-Systeme könnten besonders gefährliche oder gesundheitsgefährdende Arbeitsschritte übernehmen oder durch die Analyse der Ursachen von Unfällen und Erkrankungen die Schaffung gezielter Gegenmaßnahmen fördern (World Economic Forum 2024). Andererseits könnte KI auch Arbeitsbedingungen verschlechtern, menschliche Fachkräfte überflüssig machen, **bestehende Ungleichheiten** verschärfen oder die Versorgung von Patientinnen und Patienten negativ beeinträchtigen. Ungeachtet dieser Befürchtungen schreitet die Integration von KI wie in vielen gesellschaftlichen Domänen auch im Bereich des Betrieblichen Gesundheitsmanagements (BGM) rapide voran.

Dieser Beitrag beleuchtet **ethische Aspekte** von KI im BGM. Nach einer kurzen Einführung von KI-Ethik als Forschungsfeld illustrieren die Unterkapitel die Bedeutung und Relevanz verschiedener ethischer Dimensionen (Fairness, Privatheit, Transparenz, menschliche Aufsicht und Letztentscheidung<sup>1</sup>) für die Entwicklung ethisch verträglicher KI-Systeme. Die verschiedenen Werte werden dafür mit beispielhaften KI-Anwendungen im BGM in Bezug gesetzt. Der Beitrag endet mit einer kurzen Reflexion.

### 2.2 KI-Ethik als Forschungsfeld

Die Ethik der künstlichen Intelligenz ist als ein Praxisbereich der angewandten Ethik der normativen Ethik zuzuordnen. Als wissenschaftliche Disziplin beschäftigt sich die normative Ethik mit der Begründung, Kritik oder Rechtfertigung menschlicher Verhaltensweisen und -normen (Werner 2021). Das allge-

<sup>1</sup> Die Liste der hier aufgeführten ethischen Aspekte von KI ist nicht erschöpfend; weitere Dimensionen wie z. B. Nachhaltigkeit und Accountability sind für die Entwicklung sozial verträglicher Systeme ebenfalls zu beachten.

meine Ziel der KI-Ethik ist es, den gesellschaftlichen Nutzen von KI zu maximieren, während Risiken möglichst minimiert werden. Dafür wird analysiert, wie KI-Technologie in **gesellschaftliche Dynamiken** eingebettet sind und welche Risiken, Probleme und auch Vorteile sie – gegenwärtig wie zukünftig – mit sich bringen. Forschende in der KI-Ethik arbeiten z. B. in integrativen Design-Prozessen (Spindler et al. 2020) mit technischen, kommerziellen und rechtlichen Partnern, um die Entwicklung ethisch akzeptabler Technologie voranzutreiben. Ansätze wie das *human-centered design* (HCD) oder das *value-sensitive design* (VSD; Friedman 1996) können dabei helfen, Nutzungsszenarien gemeinsam zu erproben, Werte-Konflikte zu identifizieren und entsprechende Risiken zu adressieren.

Die Beachtung ethischer Werte in der Technikgestaltung erleichtert auch die Einhaltung von Gesetzen. So wurde der *Artificial Intelligence Act* der Europäischen Union (**EU AI Act**, in Kraft getreten im August 2024) zur Förderung sicherer, vertrauenswürdiger und ethisch vertretbarer künstlicher Intelligenz auch auf der Grundlage ethischer Richtlinien und damit explizit wertebasiert entwickelt (HLEG 2019). Der AI Act klassifiziert KI-Anwendungen in verschiedene Risikokategorien. Anwendungen im Gesundheits- und im Arbeitsbereich sind oft als Hochrisiko-KI klassifiziert, da sie bei fehlerhafter oder missbräuchlicher Nutzung zu **schwerwiegenden Folgen** für die Betroffenen führen können. Gerade in Hochrisikobereichen sind KI-ethische Analysen und Überlegungen daher unumgänglich, um gesellschaftliche Risiken zu antizipieren und die Entwicklung sozial verträglicher Systeme zu fördern.

### 2.3 Fairness und Antidiskriminierung: Ist KI objektiv?

Menschliches Handeln kann bewusst oder unbewusst von Stereotypen und Vorurteilen beeinflusst sein, z. B. gegenüber ethnischen Min-

derheiten, Menschen mit Behinderung, Frauen oder Personen mit bestimmten religiösen oder politischen Weltanschauungen. Im Gegensatz dazu sind computergetriebene Entscheidungen vermeintlich neutral, objektiv, vorurteilsfrei und damit *fair*. Doch eine zunehmende Anzahl an Skandalen, z. B. bezüglich sexistischer (Gomez 2024), rassistischer (Ledford 2019) oder ableistischer (Wall und Shellman 2021) Diskriminierung durch KI-Systeme, deutet darauf hin, dass das nicht unbedingt den Tatsachen entspricht. *Biases* (deutsch: Verzerrungen), z. B. aufgrund voreingenommener oder fehlerhafter Daten, können die Vorhersagen von KI-Systemen so verzerrn, dass es zur Diskriminierung von Einzelpersonen oder Personengruppen kommt (Barocas et al. 2019). KI ist hochgradig abhängig von der Verfügbarkeit großer Mengen an **Trainingsdaten**, anhand derer sie „lernt“, Zusammenhänge und Muster zu erkennen. Diese Daten, z. B. Texte oder Fotos, stammen aus komplexen sozialen Kontexten und können gesellschaftliche Umstände – inklusive bestehender Diskriminierungen – widerspiegeln. Zusätzlich sind Datensätze meist menschlich vorsortiert oder gekennzeichnet bzw. annotiert, was weitere Vorurteile einspeisen kann. KI-Systeme können daher menschliche Biases mitlernen und sie im laufenden Betrieb entsprechend reproduzieren und automatisieren (Eubanks 2018). Es besteht eine Wechselbeziehung zwischen Gesellschaft und Technologie: KI-Technologien sind **sozio-technische Systeme**, die in gesellschaftliche Dynamiken eingebettet sind und diese wiederum beeinflussen. Da KI-Systeme also grundsätzlich zu einer Automatisierung gesellschaftlicher Ungleichheiten führen können, tangieren sie Fragen sozialer Gerechtigkeit, die Schutzmaßnahmen zur Gleichbehandlung und Anti-Diskriminierung fordern.

Statistische Methoden der KI-Fairness sollen soziale Diskriminierung durch KI-Modelle abmildern oder beseitigen (vgl. den Beitrag von Adolph in diesem Band). Häufig kommen Fairness-Metriken zum Einsatz, anhand derer mögliche Verzerrungen (englisch: *Biases*) gemessen und Modelle nach bestimmten Vorstel-

lungen von Fairness optimiert werden können. Wird z. B. die häufig genutzte Metrik der demographischen Parität angewandt, wird das System darauf untersucht, ob es für verschiedene Personengruppen die gleichen Vorhersagen trifft (Mehrabi et al. 2021). Trainierte Modelle werden idealerweise vor der Markteinführung des Systems ausgiebig auf Verzerrungen getestet und gegebenenfalls anhand weiterer Datensätze neu trainiert. Verzerrungen können auch im laufenden Betrieb auftauchen, z. B. wenn Personen korrekte Modellvorhersagen aufgrund eigener Vorurteile wiederholt ablehnen und das System diese Entscheidungen mitlernt. KI-Systeme müssen daher regelmäßig auf diskriminierende Tendenzen untersucht und an dynamische gesetzliche, ethische und technische Standards angepasst werden. Zusätzlich hängt die Identifizierung geeigneter Fairness-Verfahren über technische Fragen hinaus von unterschiedlichen Gerechtigkeitsvorstellungen und Wertesystemen ab; Fairness ist in hohem Maße **situations- und kontextspezifisch** (Hagendorff 2019; Wachter et al. 2020). Vorherrschende Stereotype und Vorurteile sowie deren Folgen können sich z. B. je nach kulturellem Umfeld oder geographischer Lage grundlegend unterscheiden. Abseits mathematischer Modelle müssen realweltliche soziale Umstände, historische und strukturelle Macht-dynamiken berücksichtigt werden, um besonders vulnerable Personengruppen und potenzielle Diskriminierungen zu identifizieren (John-Mathews et al. 2022). Rein statistisch verstandene Fairness kann Diskriminierung im realen Betrieb von KI-Systemen also nicht verhindern. KI-Fairness muss ganzheitlich, aus inter- und transdisziplinärer Sicht von den beteiligten Stakeholdern verhandelt werden (Brandner und Hirsbrunner 2023). Das bedeutet, dass in die Technikentwicklung sowohl Expertisen aus verschiedenen Disziplinen wie Informatik, Ethik, Rechts- und Sozialwissenschaften, aber auch außerwissenschaftliche gesellschaftliche Perspektiven, wie die Meinungen und Bedarfe von Bürgern und Bürgerinnen, miteinbezogen werden sollten.

## ■ ■ Beispiel: KI für medizinische Diagnostik und Vorsorge

KI-Technologien sind vielversprechende medizinische Diagnose-Tools (Khalifa und Alabdawiy 2024). Im BGM könnten sie zur Früherkennung einer Vielzahl von – auch arbeitsbedingten – Erkrankungen wie Herzkreislauf-erkrankungen, Lungenleiden oder bestimmter Krebsarten eingesetzt werden, um zeitige **Interventions- und Präventionsstrategien** zu ermöglichen. Auch könnte KI anhand von Krankheitsverläufen und Arbeitsanforderungen die Dauer von Arbeitsunfähigkeiten abschätzen und bei der Erstellung von Rehabilitationsplänen assistieren. Doch gerade in diesem sensiblen Bereich zeigt sich auch das Diskriminierungspotenzial von KI durch ungeeignete, nicht repräsentative oder schlicht mangelhafte Trainingsdaten (Brandner et al. 2023). Wenn z. B. eine KI für Hautkrebs-Screening anhand von Datensätzen lernt, die hauptsächlich Beispieldaten von hellhäutigen Personen enthalten, ist das resultierende System möglicherweise nicht in der Lage, verdächtige Hautstellen auf dunklerer Haut präzise zu identifizieren (Davis 2021). Ein anderes Beispiel ist geschlechtsbezogene Diskriminierung, da KI-Tools manche Krankheiten bei Männern verlässlicher diagnostizieren können als bei Frauen (Straw und Wu 2022). KI-Modelle können auch sogenannte Stellvertreter-Indikatoren mitlernen, die trotz des Weglassens geschützter Merkmale auf diese schließen lassen (Datta et al. 2017). So assoziierte ein algorithmisches System in den USA, das zur Risikoklassifizierung von Patientinnen und Patienten eingesetzt wurde, niedrigere Gesundheitsausgaben mit einem geringeren gesundheitlichen Risiko; Personen, denen ein hohes Risiko zugeschrieben wurde, erhielten zusätzliche Behandlungen. Eine Analyse ergab allerdings, dass das System damit zur Diskriminierung schwarzer Personen beitrug, da diese aufgrund eines allgemein schlechteren Zugangs zu medizinischen Ressourcen weniger Ausgaben verursachen, nicht wegen eines niedrigeren Erkrankungsgrads (Ledford 2019).

Das Risiko der Diskriminierung ist nicht auf Diagnose-Tools beschränkt. Auch bei anderen KI-Anwendungen im Arbeitsbereich wie Systemen, die bei Entscheidungen über Beschäftigte unterstützen sollen, können Fairness-Probleme bestehen. Beispielsweise können Systeme zur KI-gestützten Personalauswahl oder Personalförderung Menschen mit Behinderung, sprachlichen Beeinträchtigungen, Neurodivergenzen und anderen gesundheitlichen Beeinträchtigungen diskriminieren (Buyl et al. 2022; Kim und Bodie 2021).

Auch beim Einsatz von KI im BGM kann es daher zur **Diskriminierung von Beschäftigten** kommen. Ein KI-System könnte z. B. dazu eingesetzt werden, die Dauer der Entgeltfortzahlung bei bestimmten Krankheiten festzulegen. Beruht die Kalkulation auf der durchschnittlichen Dauer von beruflichen Fehlzeiten, mag diese bei durchschnittlichen Beschäftigten akkurat sein; werden jedoch Faktoren wie chronische Erkrankungen, Behinderungen oder andere, individuelle Gesundheitsfaktoren nicht miteinbezogen, kann dies schnell zu einer unfairen Behandlung von Personen führen, die länger arbeitsunfähig bleiben. Einerseits muss daher darüber hinaus, wie akkurat ein KI-Tool insgesamt ist, auch bedacht werden, für *wen* es akkurat ist und für *wen* möglicherweise nicht. Alle potenziellen Zielpopulationen sollten im Voraus analysiert werden und es sollte in späteren Phasen erneut überprüft werden, ob im Einsatz weitere relevante Gruppen dazukommen. Andererseits muss gerade in besonders sensiblen Kontexten wie Krankheitszeiten und Entgeltfortzahlung, die großen Einfluss auf das gesundheitliche und finanzielle Wohlergehen von Beschäftigten haben können, kritisch hinterfragt werden, ob der Einsatz von KI-Systemen allgemein vertretbar und förderlich ist.

## 2.4 Privatheit versus Transparenz: Werte im Konflikt?

Ethische Werte können in einem **Spannungsverhältnis** stehen, was sich illustrativ am Verhältnis von Privatheit und Transparenz verdeutlichen lässt. Transparenz bezieht sich auf die Problematik, dass KI-Systeme häufig eine opake „Black-Box“-Architektur haben (Burrell 2016); das bedeutet, dass KI-Prozesse oft hochkomplex und für die menschliche Kognition – vor allem für Personen ohne die entsprechende Expertise – kaum fass- und interpretierbar sind. Die Förderung von Transparenz durch die Offenlegung und Erklärung von KI-Funktionsweisen, Datensätzen, Algorithmen und Modellen (Turilli und Floridi 2009) kann aber gleichzeitig zu Privatheitsproblemen führen (Keller 2022). Privatheit betrifft den Ausschluss verschiedener Personen oder Gruppen von der Kenntnis bestimmter Aspekte oder Daten aus dem Leben einer Person. Sie dient den wesentlichen Interessen der Selbstentfaltung, Selbstverwirklichung und der freien Bildung sozialer Beziehungen. Ihr Schutz ist erforderlich für die Wahrung individueller Autonomie und Handlungsfähigkeit (Rachels 1975). Über den Schutz des Individuums hinaus ist Privatheit auch ein für die kollektive Wohlfahrt notwendiges Gemeingut (Heesen et al. 2022). In datenschutzsensiblen Kontexten wie dem Gesundheitswesen ist vor allem das Recht auf **informationelle Selbstbestimmung** relevant, das Individuen vor der unkontrollierbaren Freigabe personenbezogener Informationen<sup>2</sup> schützen soll. Kompromisse zwischen Transparenz und Datenschutzbedenken müssen also im jeweiligen Kontext der Anwendung bedacht werden. Andererseits kann mangelnde Transparenz nicht mit Datenschutzbedenken gerechtfertigt werden: So lässt sich die rechtlich vorgeschriebene ausdrückliche, **freiwillige Einwilligung**

<sup>2</sup> D. h. Informationen wie Name oder Geburtsdatum, die sich eindeutig einer Person zuordnen lassen, aber auch personenbeziehbare sachliche Daten wie Konsumverhalten oder Aufenthaltsorte (Heesen 2017).

(englisch: *informed consent*) zur Verarbeitung personenbezogener Daten nur dann einholen, wenn der einwilligenden Person gegenüber ausreichend Transparenz besteht und sie in der Lage ist, die Tragweite ihrer Entscheidung zu verstehen. Die Nutzenden eines KI-Systems müssen also explizit und in verständlicher Sprache darüber informiert werden, dass, wie und zu welchem Zweck ihre Daten durch ein KI-System verarbeitet werden sowie welche Widerrufs- und Löschmöglichkeiten es gibt. Individuelle **Data Literacy** (deutsch: *Datenkompetenz*) stößt beim Versuch, das Zusammenspiel verschiedener, immer komplexer werdender Technologien im Detail nachzuvollziehen, schnell an ihre Grenzen (Heesen 2017). Ein bekannter Ansatz für privatheitsschützende Technikgestaltung, der auch von der Datenschutz-Grundverordnung (DSGVO) aufgegriffen wird, ist der Datenschutz durch Technikgestaltung (englisch: *privacy by design*; Cavoukian 2009). Dieser Ansatz weist beteiligten Unternehmen Verantwortung über die Erfüllung gesetzlicher Mindeststandards hinaus zu:

» „*Datenschutz muss standardmäßig in vernetzte Datensysteme und Technologien integriert werden. Datenschutz muss ein integraler Bestandteil organisatorischer Prioritäten, Projektziele, Designprozesse und Planungsabläufe werden.*“<sup>3</sup> (Cavoukian 2009, S. 1)

Privatheitsschützende Technikgestaltung erfordert also eine Infrastruktur aus normativen Rahmenbedingungen sowie technischen und organisatorischen Maßnahmen (Heesen et al. 2022). Z. B. sollte für ein KI-System, das sensible Gesundheitsdaten von Beschäftigten verarbeitet, auf organisatorischer Seite klar festgelegt und dokumentiert werden, wer wann Zugriff auf die Daten hat, und die entsprechenden Personen sollten durch Datenschutzschulungen sensibilisiert sein. Auf technischer Ebene können z. B. vorgeschriebene Lösch-

fristen für personenbezogene Daten in die Technologie selbst implementiert sein.

### ■ ■ Beispiel: Tracking von Beschäftigten

Die individuelle Erfassung gesundheitsrelevanter Daten wie der täglichen Schrittzahl oder des Blutdrucks ist im Privatleben vieler Personen schon lange verankert (Reichert 2015). Entsprechende Technologien, z. B. Smartphone-Apps und intelligente tragbare Geräte oder Sensoren (englisch: *smart wearables*), werden vermehrt auch für den Gebrauch im Arbeitsleben entwickelt, z. B. um Arbeitsunfälle zu verhindern, gesundheitliche Probleme frühzeitig zu erkennen und individuelle Maßnahmen der Gesundheitsförderung zu empfehlen. Die Analyse von Gesundheitsdaten ermöglicht tiefere Einblicke in den körperlichen Zustand und den privaten Lebenswandel einer Person; so wird beispielsweise die Messung von Ermüdung, Stress oder auch des Alkohol- und Drogenkonsums von Beschäftigten erforscht (El-Helaly 2024; Moshawrab et al. 2022). Bei dieser Art sensibler gesundheitsrelevanter Daten besteht großes **Missbrauchspotenzial** (Deutscher Ethikrat 2017).

Nicht datenschutzkonformes Tracking von Beschäftigten kann teure Folgen für Unternehmen haben: 2024 wurde die Handelsfirma Amazon in Frankreich mit einer Geldstrafe in Millionenhöhe belegt, da deren Tracking der Performance und Pausen von Beschäftigten gegen die DSGVO verstieß. Doch auch über rechtliche Fragen hinaus sollten Unternehmen die Entscheidungsfreiheit von Arbeitnehmenden bezüglich der eigenen Gesundheit ernst nehmen; gerade eine Kontrolle, die über eine mögliche Gefährdung am Arbeitsplatz hinaus in den persönlichen Lebenswandel eingreift, kann ein Gefühl von **Überwachung** auslösen. Haben Personen das Gefühl, überwacht zu werden, kann dies zu mehr Selbstdisziplinierung führen, auch als *chilling effects* bezeichnet (Penney 2021). Angenommen, die Belegschaft eines Unternehmens wird darüber informiert, dass zur Erfassung von Müdigkeit ein System implementiert wurde, dass während der Arbeit am Bildschirm Augenbewegungen

<sup>3</sup> Übersetzung aus dem Englischen durch die Autorin.

analysiert. Das System ist für die Beschäftigten nicht wahrnehmbar, sodass sie nicht wissen, wann es aktiv ist, wie es genau funktioniert, ob und wann es Müdigkeit bei einzelnen Personen feststellt. Zusätzlich fragen sie sich, ob das System wirklich nur Ermüdung erfasst oder auch ihre Leistung überwachen und bewerten soll. Aus Sorge vor negativen Konsequenzen beginnen die Beschäftigten daher, ihr Verhalten an die vermutete Funktionsweise des Systems anzupassen, versuchen z. B. weniger zu blinzeln, verwenden Augentropfen, erhöhen ihren Koffeinkonsum und schränken Freizeitaktivitäten ein, um möglichst keinerlei Zeichen von Ermüdung während der Arbeit zu zeigen.

Im Gegensatz zu diesem Negativbeispiel kann transparente Kommunikation darüber, welche Daten genau analysiert werden, zu welchem Zweck und mit welchen potenziellen Folgen, das Vertrauen und die Akzeptanz der Belegschaft positiv beeinflussen (BSI 2024). Um die Bedingungen für eine freiwillige Einwilligung zu erfüllen und die informationelle Selbstbestimmung zu stärken, sollten auch realistische „Opt-out“-Möglichkeiten geboten werden, also die Option, die Zustimmung zur Datenverarbeitung zu verweigern, ohne dass daraus Nachteile entstehen. Das deutsche Bundesdatenschutzgesetz (BDSG) weist bezüglich der Freiwilligkeit auf die **hierarchische Beziehung** zwischen Beschäftigten und Unternehmen hin, da Beschäftigte häufig auf ihren Arbeitsplatz angewiesen und damit vulnerabel sind. Ohne „Opt-out“-Option können sie sich schnell gezwungen sehen, ihre Privatheit zugunsten der beruflichen Sicherheit aufzugeben (Behrendt und Loh 2022), womit die Bedingungen für eine freiwillige Einwilligung nicht erfüllt wären.

## 2.5 Menschliche Autonomie – Aufsicht und Letztentscheidung

Viele KI-Horrorszenarien beziehen sich auf einen totalen menschlichen Kontrollverlust über Technologie. Gerade mediale Darstellungen konzentrieren sich häufig auf katastrophale Szenarien wie rebellierende humanoide Roboter oder eine algorithmisch gesteuerte Welt ohne Selbstbestimmung. Doch auch abseits von Dystopien kann ein Mangel an menschlicher Kontrolle beim Einsatz von KI-Systemen schwerwiegende Folgen haben; beispielsweise können ohne effektive menschliche Aufsicht Verzerrungen länger unerkannt bleiben, was wie weiter oben diskutiert zu Diskriminierung führen kann. Auch die oft befürchtete **Automatisierung von Arbeitsplätzen** kann durch das Fehlen von Mechanismen zur menschlichen Aufsicht und Letztentscheidung verschärft werden.

Der AI Act verpflichtet Bereitsteller und Anbieter von Hochrisiko-KI, eine menschliche Aufsicht über ihre Systeme zu gewährleisten. Das bedeutet, dass natürliche Personen die Funktionsweise von KI-Systemen überwachen und sicherstellen, dass sie bestimmungsgemäß verwendet werden und verlässliche Ergebnisse liefern. Die Verantwortung darf dabei nicht allein bei Individuen wie Endnutzenden liegen; die Systeme selbst müssen über Mechanismen verfügen, die menschliche Aufsichtspersonen anleiten und informieren, damit diese entscheiden können, ob, wann und wie sie in den Prozess eingreifen oder das System abschalten müssen, wenn es nicht wie vorgesehen funktioniert. Dazu gehört auch die von der DSGVO vorgeschriebene menschliche Letztentscheidung, die das grundlegende Recht sichern soll, nicht allein algorithmischen Entscheidungen unterworfen zu sein. Dementsprechend dürfen Entscheidungen nicht allein bei KI-Systemen liegen; die Entscheidung von Nutzenden soll immer Vorrang vor der Entscheidung des Systems haben. So dürfen auch im Kontext des BGM keine vollautomati-

tisierten Entscheidungen getroffen werden, die signifikante Auswirkungen auf das (Arbeits-)Leben von Beschäftigten haben; Aufsichtspersonen müssen die Systemprozesse zu jedem Zeitpunkt nachvollziehen und Ergebnisse kritisch hinterfragen und einordnen können.

Die Verpflichtung zu menschlicher Aufsicht und Letztentscheidung reflektiert die Bedeutsamkeit ethischer Gesichtspunkte für die Regulierung von KI, da sie die menschliche Handlungsfähigkeit und Autonomie in den Fokus rückt. Ohne kompetente menschliche Aufsicht und Letztentscheidung kann die Komplexität und wahrgenommene Autorität von KI-gestützten Prozessen und Outputs dazu führen, dass fehlerhafte Entscheidungen und Verzerrungen nicht erkannt werden. Wenn KI-Entscheidungen blind akzeptiert werden, kann es zu einer **Autoritätsumkehr** kommen zwischen Nutzenden und Systemen, die eigentlich der menschlichen Entscheidungsfindung nur assistieren sollen (Yeung et al. 2021). Gleichzeitig ist die konkrete Ausgestaltung effizienter menschlicher Aufsicht unklar, da das Konzept oft vage und lückenhaft bleibt (Enqvist 2023). Um KI-Risiken effektiv erkennen und mindern zu können, müssen Aufsichtspersonen u. a. die nötigen Kompetenzen, Autoritäten, Entscheidungsbefugnisse, Supportstrukturen sowie angemessene Intentionen haben (Sterz et al. 2024). Menschliche Aufsicht ist also ein wesentlicher Aspekt ethisch akzeptabler Technologien, kann aber wiederum nicht als Allheilmittel für ethische Probleme von KI-Technologien gelten – vor allem nicht, wenn sie als individuelle Aufgabe statt als komplexe Infrastruktur verschiedener Verantwortlichkeiten verstanden wird (Zweig et al. 2018).

Weiterhin hängt menschliche Kontrolle auch mit anderen ethischen Dimensionen zusammen: KI-Nutzende müssen durch technologische Transparenz dazu befähigt werden, die Vertrauenswürdigkeit von Modellen, Daten und anderen Elementen kritisch hinterfragen, überprüfen und korrigieren zu können. Wissenschaftliche Ansätze wie reflexives Design (Sengers et al. 2005) und reflexive Informatik (Baumer 2015) haben aufgezeigt, wie Irrita-

tionen und Angebote zu Reflexion und **kritischer Evaluation** von Technologie aktiv in Programme integriert werden können, um die menschliche Kontrolle zu fördern. Ein Beispiel für ein solches Angebot wäre ein offenes Textfeld, in dem Nutzende eigene Gedanken und Einschätzungen zu dem Ergebnis eines KI-Systems eintragen können. Maßnahmen wie diese können zu einem reflektierten Umgang mit KI führen, da sie Personen aktiv zur tiefergehenden Beschäftigung mit der Technologie auffordern und ihnen gegebenenfalls aufzeigen, welche Systemprozesse nicht transparent genug sind.

### ■ ■ Beispiel: Psychische Gesundheit am Arbeitsplatz

Auch die psychische Gesundheit von Beschäftigten kann durch KI-Tools potenziell gemessen und verbessert werden. Z. B. könnte KI psychische Belastung (Doki et al. 2021) bis hin zu akuten Krisen (Swaminathan et al. 2023) von Beschäftigten feststellen und Gegenmaßnahmen einleiten. Durch frühzeitige Detektion von erhöhtem Stress könnten Burnouts verhindert werden. Allerdings sind KI-Tools zur **Identifikation affektiver Zustände** – wie Depression, Angst oder Stress – besonders kontrovers. Da emotionale Ausdrücke von individuellen sowie kulturellen und sozialen Faktoren abhängen können (Feldman Barrett et al. 2019), laufen diese Systeme besonders Gefahr, schädliche Stereotype zu perpetuieren und durch unzuverlässige Ergebnisse diskriminierend zu wirken. Sie können stigmatisierend wirken (Corvite et al. 2023) oder bestehende Ungleichheiten verfestigen (Manokha 2019). Empirische Untersuchungen haben z. B. gezeigt, dass Emotionserkennungssysteme dazu neigen, Gesichtern von schwarzen Personen mehr negative Emotionen zuzuweisen (Rhue 2018). Das unterstreicht, dass KI keineswegs selbstständig Entscheidungen treffen sollte, die reale Auswirkungen auf die – psychische wie physische – Gesundheit von Beschäftigten haben können. Gleichzeitig stellt sich die Frage, wie menschliche Aufsicht in solchen Fällen effektiv ausgestaltet werden kann. Identifiziert

bspw. eine KI eine akute psychische Krise bei einem oder einer Beschäftigten, bleibt keine Zeit, die Entscheidungsfindung des Systems im Detail nachzuvollziehen und zu überprüfen. Dass das System womöglich falsch liegt, muss bei jeder Intervention in Kauf genommen werden.

Auch allgemein kann der Einsatz von KI das psychische Wohlergehen von Beschäftigten negativ beeinträchtigen. So stellt gerade die Sorge, von einer KI ersetzt zu werden, einen Stressfaktor am Arbeitsplatz dar (El-Helaly 2024) und Beschäftigte können den vermehrten Einsatz von KI als **Autonomie- und Kontrollverlust** erleben (Niehaus et al. 2022). Dieses Empfinden kann die Arbeitsmotivation verringern und wiederum das Risiko eines Burnouts erhöhen (World Economic Forum 2024). Die Implementierung von Strategien zur menschlichen Aufsicht und Letzentscheidung stellt also auch ein Mittel dar, die psychische Gesundheit von Beschäftigten zu verbessern. Dafür sollten verschiedene Verantwortlichkeiten und Aufgaben bezüglich des KI-Einsatzes klar definiert und kommuniziert werden. Beschäftigte sollten die Möglichkeit haben, Feedback zu System-Entscheidungen zu geben, um die kontinuierliche Verbesserung und Anpassung an reale Anforderungen zu fördern. Menschliche Aufsichtspersonen sollten in erforderlichen Schulungen auch für spezifische ethische Risiken und Problematiken des jeweiligen Systems, z. B. in Bezug auf Fairness, Privatheit und Transparenz, sensibilisiert werden.

## 2.6 Fazit

KI-Technologien haben das Potenzial, durch gezielte Präventions- und Interventionsmaßnahmen, Strategien zur Effizienzsteigerung oder die Übernahme riskanter Aufgaben die Gesundheit, Sicherheit und Produktivität von Beschäftigten zu verbessern. Doch es bleibt zu betonen, dass ihre Ergebnisse auf mathematischen Wahrscheinlichkeiten beruhen – das bedeutet nicht, dass sie garantiert richtig sind.

Dennoch werden sie z. B. als Risikoklassifizierungen auf reale soziale Situationen angewandt, was – wie in diesem Beitrag dargestellt – große Gefahren für das individuelle und kollektive Wohlergehen birgt. Besonders wenn Fehlerwahrscheinlichkeiten sehr hoch sind, die technologische Entwicklung noch nicht ausgereift ist oder menschliche Aufsicht nicht gewährleistet werden kann, können die Risiken die Vorteile von KI überwiegen (Lischka und Klingel 2017). Ob und wie KI-Systeme auf ethisch vertretbare Weise eingesetzt werden können, ist dabei stark kontextspezifisch.

Vor jeder Anwendung in sensiblen Bereichen wie dem BGM müssen Systeme daher auf die Verlässlichkeit ihrer Ergebnisse, die Qualität und Eignung ihrer Trainingsdaten sowie die beabsichtigten Einsatzzwecke und Zielgruppen überprüft werden. Während der Entwicklung können die in diesem Beitrag genannten Methoden von integrativem Design, HCD oder VSD dabei helfen, ethische Prinzipien von Beginn an mitzudenken und in die Technologie zu integrieren. Doch selbst wenn die grundsätzliche Eignung eines Systems festgestellt wurde, sollten die konkrete Implementierung und Ausgestaltung von KI in der Praxis nicht über die Köpfe von Beschäftigten hinweg beschlossen, sondern mit allen betroffenen Personen gemeinsam ausgehandelt werden. Es bedarf einer kontinuierlichen institutionellen Auseinandersetzung mit ethischen Aspekten von KI, z. B. durch die Schaffung von Strukturen für effektive, informierte menschliche Aufsicht, die transparente Festlegung und Dokumentation von Verantwortlichkeiten, Ressourcen für Unterstützung, Kritik und Feedback bezüglich der Technologie sowie durch Schulungen, die alle Beteiligten für Problematiken wie Fairness, Transparenz und Privatheit sensibilisieren.

## Literatur

Barocas S, Hardt M, Narayanan A (2019) Fairness and machine learning: limitations and opportunities. <https://fairmlbook.org/>. Zugegriffen: 2. März 2025

- 2**
- Baumer E (2015) Reflective Informatics: conceptual dimensions for designing technologies of reflection. Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems, S 585–594 <https://doi.org/10.1145/2702123.2702234>
- Behrendt H, Loh W (2022) Informed consent and algorithmic discrimination – is giving away your data the new vulnerable? Rev Soc Econ 80(1):58–84. <https://doi.org/10.1080/00346764.2022.2027506>
- Brandner LT, Hirsbrunner SD (2023) Algorithmische Fairness in der polizeilichen Ermittlungsarbeit: Ethische Analyse von Verfahren des maschinellen Lernens zur Gesichtserkennung. TATuP 32(1):24–29. <https://doi.org/10.14512/tatup.32.1.24>
- Brandner LT, Mahlow P, Wilken A et al (2023) How data quality determines AI fairness: the case of automated interviewing. Proc. 2nd European Workshop on Algorithmic Fairness, Winterthur (<https://ceur-ws.org/Vol-3442/>)
- BSI – Bundesamt für Sicherheit in der Informationstechnik (2024) Transparenz von KI-Systemen. <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Whitepaper-Transparenz-KI-Systeme.pdf>. Zugegriffen: 2. März 2025
- Burrell J (2016) How the machine ‘thinks’: understanding opacity in machine learning algorithms. Big Data Soc. <https://doi.org/10.1177/2053951715622512>
- Buyl M, Cociancig C, Frattone C et al (2022) Tackling algorithmic disability discrimination in the hiring process: an ethical, legal and technical analysis. In: Proc. 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’22), S 1071–1082 <https://doi.org/10.1145/3531146.3533169>
- Cancer Research Institute (2024) AI and cancer: the emerging revolution. <https://www.cancerresearch.org/blog/january-2025/ai-cancer>. Zugegriffen: 7. März 2025
- Cavoukian A (2009) Privacy by design: the 7 foundational principles. <https://privacy.ucsc.edu/resources/privacy-by-design--foundational-principles.pdf>. Zugegriffen: 1. März 2025
- Corvite S, Roemmich K, Rosenberg T et al (2023) Data subjects’ perspectives on emotion artificial intelligence use in the workplace: a relational ethics. Lens Proc Acm Hum-comput Interact 7:CSCW1. <https://doi.org/10.1145/3579600>
- Datta A et al (2017) Proxy non-discrimination in data-driven systems <https://doi.org/10.48550/ARXIV.1707.08120>
- Davis N (2021) AI skin cancer diagnoses risk being less accurate for dark skin – study. <https://www.theguardian.com/society/2021/nov/09/ai-skin-cancer-diagnoses-risk-being-less-accurate-for-dark-skin-study>. Zugegriffen: 7. März 2025
- Deutscher Ethikrat (2017) Big Data und Gesundheit: Datensouveränität als informationelle Freiheitsgestaltung Stellungnahme. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf>. Zugegriffen: 1. März 2025
- Doki S et al (2021) Comparison of predicted psychological distress among workers between artificial intelligence and psychiatrists: a cross-sectional study in Tsukuba Science City. <https://bmjopen.bmjjournals.com/content/11/6/e046265>. Zugegriffen: 1. März 2025
- EI-Helaly M (2024) Artificial intelligence and occupational health and safety, benefits and drawbacks. Med Lav. <https://doi.org/10.23749/mdl.v115i2.15835>
- Enqvist L (2023) Human oversight’ in the EU artificial intelligence act: what, when and by whom? Law Innov Technol 15(2):508–535. <https://doi.org/10.1080/17579961.2023.2245683>
- Eubanks V (2018) Automating inequality. How high-tech tools profile, police, and punish the poor. St. Martin’s Press, New York
- Feldman Barrett L, Adolphs R, Marsella S et al (2019) Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. Psychol Sci Public Interest 20(1):1–68. <https://doi.org/10.1177/1529100619832930>
- Friedman B (1996) Value-sensitive design. Interactions 3(6):16–23
- Gomez C (2024) KI auf Sexismus trainiert? So diskriminiert die Technik. <https://www.fr.de/verbraucher/sexismus-rassismus-ki-diskriminierung-unsichtbar-benachteiligung-frauen-forschung-92861297.html>. Zugegriffen: 1. März 2025
- Hagendorff T (2019) Maschinelles Lernen und Diskriminierung. Probleme und Lösungsansätze. Österreich Z Soziol 44(S1):53–66. <https://doi.org/10.1007/s11614-019-00347-2>
- Heesen J (2017) Informationelle Selbstbestimmung (Grundbegriffe der Kommunikations- und Medienethik). Commun Social 50(4):495–500. <https://doi.org/10.5771/0010-3497-2017-4-495>
- Heesen J, Ammich QR, Baur A et al (2022) Privatheit, Ethik und demokratische Selbstregulierung in einer digitalen Gesellschaft. In: Roßnagel A, Friedewald M (Hrsg) Die Zukunft von Privatheit und Selbstbestimmung: Analysen und Empfehlungen zum Schutz der Grundrechte in der digitalen Welt. Springer, Wiesbaden, S 161–187 [https://doi.org/10.1007/978-3-658-35263-9\\_5](https://doi.org/10.1007/978-3-658-35263-9_5)
- HLEG (2019) Ethics Guidelines for Trustworthy AI. Europäische Kommission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Zugegriffen: 1. März 2025
- ILO – Internationale Arbeitsorganisation (2023) A call for safer and healthier working environments. <https://www.ilo.org/publications/call-safer-and-healthier-working-environments>. Zugegriffen: 1. März 2025
- John-Mathews J-M, Cardon D, Balagué C (2022) From reality to world. A critical perspective on AI fair-

- ness. J Bus Ethics 178:945–959. <https://doi.org/10.1007/s10551-022-05055-8>
- Keller D (2022) User privacy vs. Platform transparency: the conflicts are real and we need to talk about them. The center for Internet and society at stanford law school. <https://cyberlaw.stanford.edu/blog/2022/04/user-privacy-vs-platform-transparency-conflicts-are-real-and-we-need-talk-about-them-0>. Zugegriffen: 3. März 2025
- Khalifa M, Albadawy M (2024) AI in diagnostic imaging: revolutionising accuracy and efficiency. Comput Methods Programs Biomed Update. <https://doi.org/10.1016/j.cmpbup.2024.100146>
- Kim P, Bodie MT (2021) Artificial intelligence and the challenges of Workplace discrimination and privacy. ABA J Labor Employ Law 35:289
- Ledford H (2019) Millions of black people affected by racial bias in health-care algorithms. Nature 574:608–609. <https://doi.org/10.1038/d41586-019-03228-6>
- Lischka K, Klingel A (2017) Wenn Maschinen Menschen bewerten. Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung. Bertelsmann (Arbeitspapier)
- Manokha I (2019) Facial analysis AI is being used in job interviews – it will probably reinforce inequality. [https://theconversation.com/facial-analysis-ai-is-being-used-in-job-interviews-it-will-probably-reinforce-inequality-124790%22%20/t%20%22\\_blank](https://theconversation.com/facial-analysis-ai-is-being-used-in-job-interviews-it-will-probably-reinforce-inequality-124790%22%20/t%20%22_blank). Zugegriffen: 3. März 2025
- Masterson V (2024) 9 ways AI is helping tackle climate change World Economic Forum. <https://www.weforum.org/stories/2024/02/ai-combat-climate-change/>. Zugegriffen: 2. März 2025
- Mehrabi N, Morstatter F, Saxena N et al (2021) A survey on bias and fairness in machine learning. ACM Comput Surv. <https://doi.org/10.1145/3457607>
- Moshawrab M, Adda M, Bouzouane A et al (2022) Smart wearables for the detection of occupational physical fatigue: a literature review. Sensors. <https://doi.org/10.3390/s22197472>
- Niehaus S, Hartwig M, Rosen PH et al (2022) An occupational safety and health perspective on human in control and AI. Front Artif Intell. <https://doi.org/10.3389/frai.2022.868382>
- Penney J (2021) Understanding chilling effects. Minn Law Rev 106:1451
- Rachels J (1975) Why privacy is important. Philos Public Aff 4(4):323–333
- Reichert R (2015) Digitale Selbstvermessung. Verdatung und soziale Kontrolle. Z Medienwiss 7(2):66–77
- Rhue L (2018) Racial influence on automated perceptions of emotions <https://doi.org/10.2139/ssrn.3281765>
- Sengers P, Boehner K, David S et al (2005) Reflective design. Proc. 4th decennial ACM conference on Critical computing. <https://doi.org/10.1145/1094562.1094569>
- Spindler M, Booz S, Gieseler H et al (2020) How to achieve integration? In: Gransche B, Manzeschke A (Hrsg) Das geteilte Ganze: Horizonte Integrierter Forschung für künftige Mensch-Technik-Verhältnisse. Springer, Wiesbaden, S 213–239 [https://doi.org/10.1007/978-3-658-26342-3\\_11](https://doi.org/10.1007/978-3-658-26342-3_11)
- Sterz S, Baum K, Biewer S (2024) On the quest for effectiveness in human oversight: interdisciplinary perspectives. Proc. 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), S 2495–2507 <https://doi.org/10.1145/3630106.3659051>
- Straw I, Wu H (2022) Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. BMJ Health Care Informatics. <https://doi.org/10.1136/bmjhci-2021-100457>
- Swaminathan A, Lopez I, Mar RAG et al (2023) Natural language processing system for rapid detection and intervention of mental health crisis chat messages. npj Digit Med 6(1):213. <https://doi.org/10.1038/s41746-023-00951-3>
- Turilli M, Floridi L (2009) The ethics of information transparency. Ethics Inf Technol 11(2):105–112. <https://doi.org/10.1007/s10676-009-9187-9>
- Wachter S, Mittelstadt B, Russell C (2020) Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI <https://doi.org/10.2139/ssrn.3547922>
- Wall & Shellman (2021) Wie KI bei der Kandidatenauswahl behinderte Menschen diskriminiert. <https://www.heise.de/hintergrund/Wie-KI-bei-der-Kandidatenauswahl-behinderte-Menschen-diskriminiert-6158869.html>. Zugegriffen: 1. März 2025
- Walther CC (2024) 5 ways to harness AI and end poverty forever. <https://www.forbes.com/sites/corneliawalther/2024/10/16/5-ways-to-harness-ai-and-end-poverty-forever/>. Zugegriffen: 1. März 2025
- Werner M (2021) Einführung in die Ethik. Springer, Heidelberg
- World Economic Forum (2024) Digital health and AI: ushering in a new era of employee productivity and well-being. [https://reports.weforum.org/docs/WEF\\_Digital\\_Health\\_and\\_AI\\_2024.pdf](https://reports.weforum.org/docs/WEF_Digital_Health_and_AI_2024.pdf). Zugegriffen: 1. März 2025
- Yeung D, Khan I, Kalra N et al (2021) Identifying systemic bias in the acquisition of machine learning decision aids for law enforcement applications <https://doi.org/10.7249/pea862-1>
- Zweig KA, Fischer S, Lischka K (2018) Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung. Bertelsmann Stiftung <https://doi.org/10.11586/2018006>



# Datenschutz- und KI-rechtliche Aspekte beim Einsatz Künstlicher Intelligenz in Unternehmen

*Rolf Schwartmann und Moritz Köhler*

## Inhaltsverzeichnis

- 3.1 KI und Gesundheitsdaten – 28**
  - 3.1.1 Daten im Lebenszyklus eines KI-Systems – 28
  - 3.1.2 Gesundheitsdaten und der European Health Data Space – 29
- 3.2 Datenschutzrechtliche Grundlagen des KI-Einsatzes – 30**
  - 3.2.1 Anwendbarkeit und Verantwortlichkeit – 30
  - 3.2.2 Rechtmäßigkeit der Verarbeitung – 31
  - 3.2.3 Verbot der automatisierten Entscheidung im Einzelfall – 32
- 3.3 Die Verordnung über künstliche Intelligenz – 33**
  - 3.3.1 Regelungsgegenstand und Akteure der KI-Verordnung – 33
  - 3.3.2 KI-Kompetenz – 34
  - 3.3.3 Risikobasierter Ansatz – 35
  - 3.3.4 Wechsel in die Anbieterrolle – 36
- 3.4 Zusammenfassung und Ausblick – 36**
- Literatur – 36**

## ■■ Zusammenfassung

*Seit dem 2. Februar 2025 gelten die ersten Vorgaben der europäischen KI-Verordnung. Der Rechtsakt tritt neben die geltenden allgemeinen Gesetze und enthält Vorschriften, die die Sicherheit der Entwicklung und des Betriebs von KI-Systemen gewährleisten sollen. Aufgrund der Bedeutung hochwertiger Daten für den erfolgreichen Einsatz Künstlicher Intelligenz ist neben der KI-Verordnung die bereits seit 2018 geltende Datenschutz-Grundverordnung ein zentrales Regelwerk für Unternehmen, die KI-Systeme implementieren und nutzen wollen. Zugleich ermöglicht die seit März 2025 geltende Verordnung über einen europäischen Gesundheitsdatenraum den Zugang zu bisher kaum erreichbaren Datensätzen im Gesundheitsbereich. Der vorliegende Beitrag gibt vor dem Hintergrund dieser Entwicklungen einen Überblick über die rechtlichen Vorgaben, die von Unternehmen beim Einsatz von KI-Systemen insbesondere im Gesundheitsbereich zu beachten sind.*

## 3.1 KI und Gesundheitsdaten

In der Vergangenheit wurde häufig betont, dass für die Entwicklung von KI-Systemen große Mengen an Daten erforderlich sind. Daran ist richtig, dass die **Verfügbarkeit** großer Datensets eine größere Auswahl eröffnet. Die qualitative Auswahl der im gesamten **Lebenszyklus** eines KI-Systems eingesetzten Daten bestimmt allerdings erst die Qualität des Systems und seiner Ausgaben. So hat sich das geflügelte Wort „garbage in, garbage out“ mittlerweile über die Grenzen der Informatik hinaus durchgesetzt – es besagt, dass ein Rechner mit hoher Wahrscheinlichkeit (aber nicht unbedingt) eine ungültige oder nicht aussagekräftige Ausgabe produziert, wenn die Eingabe ungültig oder nicht aussagekräftig ist. Es wird üblicherweise verwendet, um darauf hinzuweisen, dass Rechner nicht von sich aus korrekte bzw. aussagekräftige Eingaben von falschen bzw. nicht aussagekräftigen unterscheiden können. Im Folgenden werden zu-

nächst die verschiedenen Datenkategorien im Lebenszyklus eines KI-Systems benannt und anschließend dargelegt, wie das Recht der Europäischen Union (EU) speziell im Gesundheitsbereich für die Verfügbarkeit hochwertiger Daten sorgen will.

### 3.1.1 Daten im Lebenszyklus eines KI-Systems

Die europäische KI-Verordnung (KI-VO; Europäisches Parlament und Rat der Europäischen Union 2024), die in der EU für die Regulierung von KI-Systemen maßgeblich ist<sup>1</sup>, unterteilt die während des Lebenszyklus eines KI-Systems relevanten Daten in vier Kategorien (► Tab. 3.1).

Die ersten drei Kategorien sind für Unternehmen nur von Bedeutung, wenn eigene KI-Systeme entwickelt werden. Zur Verdeutlichung der Besonderheiten eines KI-Systems gegenüber klassischer Software seien das Training und die Bedeutung der Trainingsdaten im Folgenden dennoch vertiefend dargestellt. KI-Systeme lassen sich nach dem Verständnis der KI-VO von einfachen algorithmischen Systemen unterscheiden, indem die **Handlungsanweisungen** verglichen werden, die der Ausführung des jeweiligen Programms zugrunde liegen. Einfache algorithmische Systeme, die keine KI-Systeme sind, beruhen ausschließlich auf Regeln für das automatische Ausführen der einzelnen Operationen, die von einem menschlichen Programmierer definiert wurden.<sup>2</sup> Der menschliche Entwickler hat dem System also im Sinne einer eindeutigen „Wenn-Dann-Regel“ vorgegeben, wie es auf verschiedene Eingaben des Nutzers reagieren soll. Kann das System eine Eingabe keiner der definierten Regeln zuordnen, erzeugt es keine Ausgabe. Der Entwickler muss dem System deshalb genau vorgeben, wie es in verschiedenen Situationen zu reagieren hat.

<sup>1</sup> Art. 288 UAbs. 2 AEUV.

<sup>2</sup> ErwG 12 Satz 2 KI-VO.

**Tab. 3.1** Datenkategorien in der KI-VO

<b>Trainingsdaten</b>	Werden zum Trainieren eines KI-Systems verwendet, wobei dessen lernbare Parameter angepasst werden	Art. 3 Nr. 29 KI-VO
<b>Validierungsdaten</b>	Werden zur Evaluation des trainierten KI-Systems und zur Einstellung seiner nicht erlernbaren Parameter und seines Lernprozesses verwendet, um unter anderem eine Unter- oder Überanpassung zu vermeiden	Art. 3 Nr. 30 KI-VO
<b>Testdaten</b>	Werden für eine unabhängige Bewertung des KI-Systems verwendet, um die erwartete Leistung dieses Systems vor dessen Inverkehrbringen oder Inbetriebnahme zu bestätigen	Art. 3 Nr. 32 KI-VO
<b>Eingabedaten</b>	Die in ein KI-System eingespeisten oder von diesem direkt erfassten Daten, auf deren Grundlage das System eine Ausgabe hervorbringt	Art. 3 Nr. 33 KI-VO

Fehlzeiten-Report 2025

Manche Aufgaben sind zu komplex, um sie für ein **algorithmisches System** auf diese Weise in eindeutige Regeln zu übersetzen. Entweder sie sind zu umfangreich oder die Regeln des Programmablaufs sind dem menschlichen Entwickler schlicht selbst nicht bekannt. Hier setzen verschiedene Ansätze der KI-Forschung an. Dabei wird der Begriff der Künstlichen Intelligenz, insbesondere im Vergleich zwischen Recht und Informatik, aber auch in den Disziplinen selbst äußerst heterogen verstanden. Nicht alles, was nach der KI-VO unter die KI-Forschung fällt, wird von den Forschenden selbst als KI bezeichnet. Die Begriffswahl in der KI-VO lässt sich vor diesem Hintergrund mit guten Gründen kritisieren.

Der Ansatz der KI-Forschung im Sinne der KI-VO ist es, dem algorithmischen System ein Ziel und einen großen und bestenfalls hochwertigen Datensatz zur Verfügung zu stellen. Das System leitet die einzelnen Arbeitsschritte dann im Sinne der angesprochenen „Wenn-Dann-Regeln“ **selbstständig** aus den Daten ab. Dieser Prozess wird als Training eines KI-Systems bezeichnet, die verwendeten Daten als Trainingsdaten.

Sind die daran anknüpfende Validierung und die Tests abgeschlossen, wird das KI-System in Betrieb genommen. Im Betrieb ist die Qualität der Eingabedaten von entscheidender

Bedeutung. Ein KI-System zur Krebsdiagnose ist beispielsweise auf eine störungsfreie CT- bzw. MRT-Bildgebung angewiesen. Bei KI-Sprachsystemen werden die Eingabedaten als *Prompt* (deutsch: Aufforderung, Anweisung) bezeichnet und umfassen den gesamten Eingabetext des Nutzers, inklusive etwaiger Dokumente, die dem System zur besonderen Berücksichtigung zur Verfügung gestellt werden.

### 3.1.2 Gesundheitsdaten und der European Health Data Space

Die **Datenqualität** ist also sowohl in der Trainings- als auch in der Betriebsphase von entscheidender Bedeutung für die Qualität der Ergebnisse des KI-Systems. Auch mit Blick auf das Potenzial Künstlicher Intelligenz hat die EU-Kommission daher bereits im Jahr 2020 eine europäische Datenstrategie veröffentlicht, deren Ziel in der Ausschöpfung der Vorteile einer besseren Datennutzung liegt (Europäische Kommission 2020). Explizit werden darin mehrfach die angestrebten Verbesserungen für den Gesundheitsbereich angesprochen. Konkret soll demnach die stärkere und umfassendere Nutzung und Weiter-

verwendung von Gesundheitsdaten Innovatoren im Gesundheitswesen fördern.

Ein entscheidender Baustein der dargelegten europäischen **Gesundheitsdatenstrategie** ist im März 2025 in Kraft getreten: die Verordnung über den europäischen **Gesundheitsdatenraum** (Europäisches Parlament und Rat der Europäischen Union 2025). Im europäischen Gesundheitsdatenraum (engl.: European Health Data Space, kurz: EHDS) sollen natürliche Personen ihre elektronischen Gesundheitsdaten leicht kontrollieren können (Primärnutzung elektronischer Gesundheitsdaten<sup>3</sup>). Im vorliegenden Kontext besonders relevant ist aber das Ziel des EHDS, es Akteuren aus Forschung und Innovation sowie politischen Entscheidungsträgern zu ermöglichen, elektronische Gesundheitsdaten auf vertrauenswürdige und sichere Weise unter Wahrung der Privatsphäre zu nutzen (Sekundärnutzung elektronischer Gesundheitsdaten<sup>4</sup>; Schwartmann und Wasilewski 2024). Laut Erwägungsgrund (ErwG) 68 Satz 3 KI-VO soll gerade diese Sekundärnutzung von Gesundheitsdaten im EHDS „den diskriminierungsfreien Zugang zu Gesundheitsdaten und das Training von KI-Algorithmen mithilfe dieser Datensätze erleichtern“.

Um dieses Ziel zu erreichen, ist jeder Mitgliedstaat der EU gemäß Art. 55 Abs. 1 Satz 1 EHDS-VO verpflichtet, eine oder mehrere **Zugangsstellen für Gesundheitsdaten** zu benennen, die unter anderem über Anträge auf Zugang zu elektronischen Gesundheitsdaten zu entscheiden haben. Der Zugang ist unter anderem zu gewähren für das Trainieren, Testen und Bewerten von KI-Systemen.<sup>5</sup> In Deutschland ist eine entsprechende Zugangsstelle für Gesundheitsdaten beim Bundesinstitut für Arzneimittel und Medizinprodukte eingerichtet.<sup>6</sup> (Deutscher Bundestag 2024).

## 3.2 Datenschutzrechtliche Grundlagen des KI-Einsatzes

Der wachsenden Bedeutung der Datennutzung steht in der EU weiterhin ein ausgeprägtes Datenschutzrecht gegenüber. Dabei müssen sich Datennutzung und Datenschutz nicht zwangsläufig ausschließen, was auch der europäische Gesetzgeber anerkennt, indem er die Datenschutz-Grundverordnung (DS-GVO; Europäisches Parlament und Rat der Europäischen Union 2016) in zahlreichen Rechtsakten zur Datennutzung für grundsätzlich unberührt erklärt. Auch die KI-VO gilt neben der DS-GVO und geht dem Datenschutzrecht nicht vor.<sup>7</sup> Bei der Entwicklung, Implementierung und Nutzung von KI-Systemen sind daher die Regelungen der DS-GVO zu berücksichtigen.

### 3.2.1 Anwendbarkeit und Verantwortlichkeit

Ein Unternehmen, das KI-Systeme nutzt, hat die Vorgaben der DS-GVO zu beachten, wenn es dabei **personenbezogene Daten** verarbeitet. Daten weisen einen Personenbezug auf, wenn sie sich auf eine identifizierte oder identifizierbare Person beziehen.<sup>8</sup> Bei der Prüfung, ob eine Person identifizierbar ist, sind laut ErwG 26 Satz 3 DS-GVO alle Mittel zu berücksichtigen, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren. Dabei sind auch die Kosten der Identifizierung und der dafür erforderliche Zeitaufwand heranzuziehen. Daten, die auf dieser Grundlage keinen Personenbezug aufweisen, werden als anonymisierte Daten bezeichnet. Da die Verarbeitung **anonymisierter Daten** nicht den Vorgaben der DS-GVO unterfällt, ist sie für die Datennutzung von entscheidender Bedeutung (Stiftung Datenschutz 2023).

3 Art. 2 Abs. 2 lit. d EHDS-VO.

4 Art. 2 Abs. 2 lit. e EHDS-VO.

5 Art. 68 Abs. 1 lit. a i. V. m. Art. 53 Abs. 1 lit e Nr. ii EHDS-VO.

6 § 3 Abs. 1 GDNG.

7 Vgl. Art. 2 Abs. 7 Satz 2 KI-VO.

8 Art. 4 Nr. 1 DS-GVO.

Sofern verarbeitete Daten einen Personenbezug aufweisen, hängen die konkreten Pflichten des Unternehmens von seiner Rolle im Sinne der DS-GVO ab. Grundsätzlich unterscheidet die DS-GVO zwischen alleiniger **Verantwortlichkeit**, Auftragsverarbeitung und gemeinsamer Verantwortlichkeit. Verantwortlich ist die Person, die über die Zwecke und Mittel der Verarbeitung personenbezogener Daten entscheidet.<sup>9</sup> Zumindest, wenn ein KI-System über die eigenen Server des Unternehmens läuft (sog. On-Premise-Implementierung), ist von einer alleinigen Verantwortlichkeit des Unternehmens auszugehen (Schwartzmann et al. 2025). Läuft das System hingegen über eine Public bzw. eine Private Cloud, ist der Cloud-Anbieter als Auftragsverarbeiter im Sinne der DS-GVO anzusehen und unterliegt besonderen Pflichten (Schwartzmann et al. 2025). Ungeklärt ist bislang, in welchen Fällen beim Einsatz von KI-Systemen eine gemeinsame Verantwortlichkeit anzunehmen ist. Eine solche besteht grundsätzlich, wenn mehrere Verantwortliche gemeinsam über die Zwecke und Mittel der Datenverarbeitung bestimmen. Bei der einfachen Nutzung eines KI-Systems durch ein Unternehmen dürfte diese Voraussetzung nicht erfüllt sein. In diesem Fall wird es bei der Verantwortlichkeit des Unternehmens und ggf. der Auftragsverarbeitung durch den Cloud-Anbieter bleiben (Schwartzmann et al. 2025).

### 3.2.2 Rechtmäßigkeit der Verarbeitung

Den zentralen Grundsatz des Datenschutzrechts gibt Art. 5 Abs. 1 lit. a DS-GVO vor: Personenbezogene Daten müssen auf rechtmäßige Weise verarbeitet werden. Der **Grundsatz der Rechtmäßigkeit der Datenverarbeitung** wird konkretisiert in Art. 6 Abs. 1 UAbs. 1 DS-GVO. Eine Verarbeitung ist demnach nur rechtmäßig, wenn mindestens eine der in der Vorschrift benannten Bedingungen

erfüllt ist. Art. 6 Abs. 1 UAbs. 1 DS-GVO statuiert damit ein sog. Verbot mit Erlaubnisvorbehalt für die Verarbeitung personenbezogener Daten. Für die Verarbeitung personenbezogener Eingabedaten im Rahmen der Nutzung eines KI-Systems kommen grundsätzlich die Einwilligung (lit. a), die Erforderlichkeit zur Vertragserfüllung (lit. b) sowie berechtigte Interessen des Verantwortlichen (lit. f) in Betracht. Vor allem letztere werden in der Praxis als überzeugende Rechtsgrundlage für die Verarbeitung personenbezogener Daten im Kontext Künstlicher Intelligenz benannt (Keber 2024). Unternehmen können in einem Dreischritt prüfen, ob sie ein **berechtigtes Interesse** an einer Datenverarbeitung im Sinne von Art. 6 Abs. 1 UAbs. 1 lit. f DS-GVO haben (Schwartzmann et al. 2025):

1. Berechtigtes Interesse: Zunächst muss ein grundsätzlich berechtigtes Interesse an der Datenverarbeitung bestehen. Neben zweckabhängigen Interessen der Öffentlichkeit, zu denen unter anderem die Gesundheit gehört, kommen auch kommerzielle Interessen in Betracht.
2. Erforderlichkeit: Darüber hinaus muss die Verarbeitung der personenbezogenen Daten zur Verfolgung des berechtigten Interesses erforderlich sein. Das ist beispielsweise ausgeschlossen, wenn die Daten vor der Verarbeitung anonymisiert werden könnten.
3. Interessenabwägung: Schließlich dürfen die Interessen sowie die Grundrechte und Grundfreiheiten der betroffenen Person das Interesse des Verantwortlichen an der Verarbeitung nicht überwiegen. In die Interessenabwägung können Art und Umfang der Verarbeitung, der Grad der Transparenz, die ergriffenen technischen und organisatorischen Maßnahmen zum Schutz der personenbezogenen Daten oder auch die Erwartungen der betroffenen Person einzubeziehen sein.

Insbesondere beim Einsatz von KI-Chatbots etwa im Kundenservice sollten verantwortliche Unternehmen die **Zwecke der Datenver-**

<sup>9</sup> Art. 4 Nr. 7 DS-GVO.

**arbeitung** genau im Blick behalten. Eine betroffene Person mag der Datenverarbeitung zur Bearbeitung ihrer Anfrage zugestimmt haben. Das bedeutet allerdings nicht, dass die Eingabedaten durch das Unternehmen auch für künftige Werbeansprachen oder zur Verbesserung des Chatbots eingesetzt werden dürfen. Dazu ist eine gesonderte Einwilligung erforderlich, wobei insbesondere in Frage gestellt wird, ob die betroffene Person ausreichend informiert ist, bevor sie eine Einwilligung in die Datenverarbeitung zur Verbesserung eines KI-Systems gewährt (Keber 2024).

Während die berechtigten Interessen des Verantwortlichen oft als vielversprechende Rechtsgrundlage für die Verarbeitung personenbezogener Eingabedaten herangezogen werden (Keber 2024), ist bei der Verarbeitung von gesundheitlich relevanten Daten die spezielle Regelung des Art. 9 Abs. 2 DS-GVO zu beachten, die für **besondere Datenkategorien**, darunter Gesundheitsdaten, eine Rechtfertigung der Datenverarbeitung nur unter eingeschränkten Voraussetzungen anerkennt (Schwartmann und Köhler 2025). Eine Verarbeitung von Gesundheitsdaten aufgrund eines berechtigten Interesses des Verantwortlichen ist in Art. 9 Abs. 2 DS-GVO nicht vorgesehen. Allerdings kann eine entsprechende Datenverarbeitung zum Schutz lebenswichtiger Interessen der betroffenen Person<sup>10</sup> und beispielsweise zur Gesundheitsvorsorge bzw. zur medizinischen Diagnostik<sup>11</sup> erfolgen.

### 3.2.3 Verbot der automatisierten Entscheidung im Einzelfall

Von besonderer Bedeutung auch für den medizinischen Bereich ist daneben das von Art. 22 Abs. 1 DS-GVO statuierte **Verbot der automatisierten Entscheidung im Einzelfall**. Demnach hat die betroffene Person das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profil-

ing – beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt. Anders als der Wortlaut der Vorschrift nahelegt, sieht diese kein Recht vor, das von der betroffenen Person geltend gemacht werden muss, sondern sie stellt ein objektives Verbot auf, das vom Verantwortlichen auch ohne Zutun der betroffenen Person zu beachten ist (Europäischer Gerichtshof 2023). Der für die Auslegung der Vorschrift zuständige Europäische Gerichtshof (EuGH) entnimmt Art. 22 Abs. 1 DS-GVO in seiner viel beachteten Entscheidung zur SCHUFA drei Tatbestandsvoraussetzungen: Es muss eine „Entscheidung“ vorliegen, die „ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profiling – [beruht]“ und die „gegenüber [der betroffenen Person] rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt“ (Europäischer Gerichtshof 2023). Im Gesundheitsbereich ist dieses Verbot etwa im Rahmen von Triage-Entscheidungen bedeutsam.

Der Begriff der Entscheidung ist nach der jüngeren Rechtsprechung des EuGH weit auszulegen. Er umfasst nicht nur die letztendliche Auswahl zwischen zwei oder mehr Handlungsalternativen. Vielmehr weist der EuGH unter Berufung auf ErwG 71 Satz 1 DS-GVO darauf hin, dass eine Entscheidung aus mehreren handlungsabhängigen Phasen bestehen kann, die alle vom Entscheidungsbegriff des Art. 22 Abs. 1 DS-GVO umfasst sind. Anders als in der deutschsprachigen Literatur bislang überwiegend angenommen wurde (Paal 2023; Krämer 2020; Abel 2018; Taeger 2017), können daher auch Maßnahmen der **Entscheidungsvorbereitung** dem Verbot der automatisierten Entscheidung im Einzelfall unterfallen.

Wird also etwa im Rahmen der Entscheidungsfindung auf das Ergebnis eines KI-Systems zurückgegriffen, schließlich aber zumindest formell von einem Menschen entschieden, schließt das die Anwendbarkeit des Art. 22 Abs. 1 DS-GVO nicht schon im Ansatz aus. Genauer Prüfung bedarf in dieser Konstellation indes die Frage, ob von der entscheidungs-

10 Art. 9 Abs. 2 lit. c DS-GVO.

11 Art. 9 Abs. 2 lit. h DS-GVO.

vorbereitenden Maßnahme auch eine rechtliche Wirkung ausgeht oder diese die betroffene Person in ähnlicher Weise erheblich beeinträchtigt. Das ist nach der Rechtsprechung des EuGH der Fall, wenn die schlussendliche Entscheidung von der vorbereitenden Maßnahme „maßgeblich“ geleitet wird (Europäischer Gerichtshof 2023).

Leider hat es der EuGH versäumt, das zentrale **Kriterium der Maßgeblichkeit** näher zu definieren. Aus den Ausführungen des EuGH lässt sich lediglich ableiten, dass die Maßgeblichkeit zumindest dann anzunehmen ist, wenn das Ergebnis der automatisierten Entscheidungsvorbereitung „in nahezu allen Fällen“ übernommen wird (Europäischer Gerichtshof 2023). Damit bleibt unklar, welcher Grad menschlicher Einflussnahme erforderlich ist, um maßgeblichen Einfluss eines entscheidungsvorbereitenden KI-Systems auf die Entscheidung abzuwehren. Schon vor der SCHUFA-Entscheidung des EuGH war diese Frage umstritten. Sie wurde allerdings am Merkmal des „ausschließlichen Beruhens“ auf einer automatisierten Verarbeitung personenbezogener Daten geführt. Im Ergebnis wird dazu zu verlangen sein, dass eine natürliche Person die inhaltliche Hoheit über den Entscheidungsprozess behält. Wird eine natürliche Person lediglich formell zwischengeschaltet, genügt dies nicht, um das Verbot des Art. 22 Abs. 1 DS-GVO zu umgehen. Wie die **inhaltliche Hoheit** des Entscheidungsträgers sichergestellt werden kann, ist wiederum nicht abschließend geklärt. Grundsätzlich dürfte eine umfassende Plausibilitätsprüfung des Ergebnisses ausreichen. In besonders grundrechtsensiblen Fällen könnten aber eigene inhaltliche Erwägungen des Entscheidungsträgers, die das Ergebnis des KI-Systems ergänzen, erforderlich sein (Köhler 2025).

### 3.3 Die Verordnung über künstliche Intelligenz

Am 2. August 2024 ist die **Verordnung über künstliche Intelligenz** in Kraft getreten. Sie beansprucht stufenweise Geltung. Die Vorschriften über verbotene KI-Systeme und KI-Kompetenz gelten bereits seit dem 2. Februar 2025, ein Großteil der Vorschriften wird ab dem 2. August 2026 gelten. Volle Geltung beansprucht die KI-VO ab dem 2. August 2027.

Gemäß Art. 1 Abs. 1 KI-VO ist der Zweck der Verordnung, „das Funktionieren des Binnenmarkts zu verbessern und die Einführung einer auf den Menschen ausgerichteten und vertrauenswürdigen künstlichen Intelligenz (KI) zu fördern und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und die in der Charta verankerten Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, vor schädlichen Auswirkungen von KI-Systemen in der Union zu gewährleisten und die Innovation zu unterstützen“. Anders als teilweise wahrgenommen stellt die EU die Förderung Künstlicher Intelligenz in der Zielsetzung des Rechtsakts also sogar vor die Regulierung der Technologie zum Schutz bestimmter Werte und Rechte.

Im Zentrum der materiellen Regelungen steht hingegen der Schutz natürlicher Personen und der Werte der EU.

#### 3.3.1 Regelungsgegenstand und Akteure der KI-Verordnung

Um das angestrebte Schutzniveau zu gewährleisten, regelt die KI-VO das Inverkehrbringen und die Inbetriebnahme von KI-Systemen und KI-Modellen. Daneben gilt das allgemeine

**Tab. 3.2** Anbieter und Betreiber

<b>Anbieter</b>	Eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck entwickelt oder entwickeln lässt und es unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich	Art. 3 Nr. 3 KI-VO
<b>Betreiber</b>	Eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet	Art. 3 Nr. 4 KI-VO

Fehlzeiten-Report 2025

Recht. Dieses regelt insbesondere den Umgang mit Ein- und Ausgaben. Die KI-VO macht also beispielsweise Vorgaben dazu, wie ein KI-System, das in der Justiz zum Einsatz kommen soll, zu entwickeln ist und wie der Betrieb zu organisieren ist. Dagegen ergeben sich aus der KI-VO keine Vorgaben darüber, ob ein Richter Verfassungsrecht verletzt, wenn er einen Fall durch ein KI-System lösen lässt und das Ergebnis ungeprüft übernimmt (Schwartmann et al. 2024).

Die KI-VO adressiert explizit sechs verschiedene Akteure: Anbieter, Produkthersteller, Betreiber, Bevollmächtigter, Einführer und Händler. Von zentraler Bedeutung sind die Begriffe des Anbieters und des Betreibers (Tab. 3.2).

Kurz gesagt, versteht die KI-VO unter dem Begriff des Anbieters also die Stelle, die ein KI-System entwickelt und unter dem Begriff des Betreibers die Stelle, in deren Verantwortung ein KI-System eingesetzt wird.

Ein Unternehmen kann seiner Rolle als Betreiber nicht dadurch entgehen, dass es keine Lizenz für ein KI-System erwirbt, den beruflichen Einsatz auf privaten Accounts aber duldet oder gar genehmigt (Schwartmann und Zenner 2025). Der Einsatz erfolgt dann weiterhin in der Verantwortung des Unternehmens; es hat lediglich keine Kontrolle mehr darüber, in welchem Umfang und auf welche Weise die Mitarbeitenden das System einsetzen. Der Einsatz derartiger „Schatten-KI“ stellt für Unterneh-

men deshalb ein erhebliches Risiko dar, das durch eindeutige Richtlinien verringert werden kann.

### 3.3.2 KI-Kompetenz

Eine zentrale Regelung der KI-VO verlangt von Anbietern und Betreibern, dass sie Maßnahmen ergreifen, „um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind“.<sup>12</sup>

Eindeutige Vorgaben, etwa hinsichtlich einer abzuleistenden Zahl an Fortbildungsstunden, ergeben sich aus dieser Vorschrift nicht. Allerdings sollten Unternehmen, die KI-Systeme entwickeln oder selbst einsetzen, auf die Schulung ihrer Mitarbeitenden hinwirken. Angesichts der Haftungsrisiken, die bei einem verantwortungslosen Einsatz von KI-Systemen bestehen, sollte die Vermittlung von KI-Kompetenz aber auch ohne die gesetzliche

<sup>12</sup> Art. 4 KI-VO.

Pflicht auf der Agenda der Unternehmen stehen.

Es geht bei der Vermittlung von **KI-Kompetenz** zunächst nicht darum, jeden Mitarbeiter zu befähigen, die technischen Hintergründe eines KI-Systems zu kennen. Es genügt, wenn die Mitarbeitenden die Grundzüge der Technologie und ihrer Regulierung kennen:

- Was ist ein KI-System?
- Welche Nutzung von KI-Systemen ist gefahrlos möglich?
- Wo muss ich aufpassen?
- Was bedeutet „prompten“ und wie geht das?
- Wie setze ich mich mit KI-generierten Ergebnissen auseinander?
- Wie behalte ich als Mensch die Kontrolle über das Werkzeug KI?
- Was bedeutet der Einsatz von KI im beruflichen Alltag?
- Wo kann mir die Technik helfen, wo nicht?

### 3.3.3 Risikobasierter Ansatz

Neben dieser allgemeinen Vorgabe verfolgt der europäische Gesetzgeber mit der KI-VO einen **risikobasierten Regulierungsansatz**. KI-Systeme werden entsprechend ihrem Zweck in verschiedene Risikoklassen eingeteilt. ErwG 26 Satz 3 KI-VO nennt im Ergebnis vier Risikoklassen, tatsächlich lassen sich in der Regulierung drei unterscheiden: Einige KI-Praktiken hält der europäische Gesetzgeber für inakzeptabel und verbietet sie deshalb in einem abgeschlossenen Katalog.<sup>13</sup> Zu den verbotenen Systemen zählen auch Systeme zur Ableitung von Emotionen.<sup>14</sup> Insbesondere im Gesundheitsbereich ergeben sich hier Abgrenzungsschwierigkeiten. Das Verbot erfasst lediglich **Emotionserkennungssysteme**, der Einsatz zur Detektion physischer Zustände wie Schmerz oder Ermüdung aus medizinischen oder Sicherheitsgründen ist gestattet. Die Leitlinien der Kommission zur Konkretisierung

der verbotenen KI-Praktiken führen darüber hinaus aus, dass unter Emotionserkennungssystemen keine Systeme zu verstehen sind, die erkennen, ob eine Person krank ist (Europäische Kommission 2025). Die Leitlinien entfalten allerdings keine rechtliche Bindungswirkung und der Übergang zwischen Emotion und Krankheit kann insbesondere im Bereich psychischer Erkrankungen fließend sein. Da gängige Emotionserkennungssysteme auch nach Ansicht der Kommission derzeit ohnehin keine verlässlichen Ergebnisse liefern (Europäische Kommission 2025), sollten Unternehmen im Umgang mit diesen Systemen besser restriktiv vorgehen.

Von einigen KI-Systemen geht zwar ein hohes Risiko aus, ihr Nutzen überwiegt aber die festgestellten Risiken. Unter strengen Anforderungen ist ihr Inverkehrbringen und ihre Inbetriebnahme aber zumindest nach der KI-VO erlaubt. Ein KI-System ist hochriskant, wenn es als Sicherheitsbauteil eines speziell regulierten Produkts wie beispielsweise eines Medizinprodukts verwendet wird oder selbst ein solches Produkt ist<sup>15</sup> (Schwartzmann und Pottkämper 2024). Darüber hinaus führt die KI-VO in einem gesonderten Anhang III eine Reihe von Lebensbereichen auf, in denen eingesetzte KI-Systeme als **Hochrisiko-KI-Systeme** gelten. Dazu zählen etwa die Bereiche Personalmanagement, Bildung und Justiz. Alle anderen Systeme sind keiner Regulierung durch die KI-VO unterworfen. Das bei der Kommission angesiedelte Büro für Künstliche Intelligenz ist allerdings verpflichtet, Verhaltenskodizes zu erstellen, die freiwillig auf Systeme mit einem derart geringen Risiko angewendet werden können.<sup>16</sup>

Als vierte Risikoklasse nennt ErwG 26 Satz 3 KI-VO bestimmte KI-Systeme, von denen ein spezifisches Risiko ausgehen soll und für die in der Folge besondere **Transparenzpflichten** gelten. Da ein solches spezifisches Risiko aber grundsätzlich in jeder der drei genannten Risikoklassen auftreten kann, ist

13 Art. 5 Abs. 1 UAbs. 1 KI-VO.

14 Art. 5 Abs. 1 UAbs. 1 lit. f KI-VO.

15 Art. 6 Abs. 1 lit. a KI-VO.

16 Art. 95 Abs. 1 KI-VO.

die Vorstellung einer eigenen, abgrenzbaren „spezifischen Risikoklasse“ missverständlich. Unternehmen sollten unabhängig von der Prüfung, ob ihr KI-System nach den Kategorien der KI-VO verboten, hochriskant oder einfach ist, stets prüfen, ob es den besonderen Transparenzanforderungen des Art. 50 KI-VO unterliegt.

### 3.3.4 Wechsel in die Anbieterrolle

Den mit Abstand umfassendsten Pflichtenkatalog sieht die KI-VO für die Anbieter von Hochrisiko-KI-Systemen vor. Sie müssen unter anderem ein Risikomanagementsystem vorhalten<sup>17</sup>, für eine angemessene Daten-Governance sorgen<sup>18</sup> und Maßnahmen zur Sicherung eines hinreichenden Maßes an Genauigkeit, Robustheit und Cybersicherheit ergreifen<sup>19</sup>. Für ein Unternehmen ohne umfassende Expertise im Bereich der KI-Entwicklung sind diese Pflichten nicht zu stemmen.

Die Betreiber von KI-Systemen müssen deshalb darauf achten, nicht durch einen für das europäische Produktsicherheitsrecht typischen **Wechsel in die Anbieterrolle** einem Pflichtenkatalog zu unterfallen, der für sie nicht erfüllbar ist. Diese Gefahr besteht vor allem beim Einsatz von KI-Sprachsystemen, die die Eingabe offener Prompts in natürlicher Sprache ermöglichen. Denn gemäß Art. 25 Abs. 1 lit. c KI-VO gelten Betreiber als Anbieter eines Hochrisiko-KI-Systems, wenn sie die Zweckbestimmung eines KI-Systems, das nicht als hochriskant eingestuft wurde, so verändern, dass das betreffende KI-System zu einem Hochrisiko-KI-System wird. Was genau zur Änderung der Zweckbestimmung erforderlich ist, ist bislang nicht abschließend geklärt. Grundsätzlich ist aber auch eine **Zweckänderung** per einfacherem Prompt denkbar, sofern dieser Prompt ein hinreichendes Maß an Komplexität und Konkretisierung aufweist

<sup>17</sup> Art. 9 KI-VO.

<sup>18</sup> Art. 10 KI-VO.

<sup>19</sup> Art. 15 KI-VO.

(Schwartmann und Zenner 2025). Betreiber sollten deshalb insbesondere beim Einsatz von KI-Sprachsystemen in den Lebensbereichen des Anhangs III der KI-VO höchste Vorsicht walten lassen.

### 3.4 Zusammenfassung und Ausblick

Gerade im Gesundheitsbereich entfaltet der Einsatz von KI-Systemen ein beachtliches Potenzial. Entscheidender Faktor des erfolgreichen Einsatzes der Technologie ist der Bezug an hochwertigen Gesundheitsdaten und deren sichere Verarbeitung. Mit dem EHDS kann in diesem Bereich auf Fortschritte gehofft werden. Daneben sehen DS-GVO und KI-VO für Unternehmen, die KI-Systeme einzusetzen, ein ausdifferenziertes Regelwerk vor, um die Rechte der betroffenen Personen auch im digitalen Zeitalter zu schützen. Im Gesundheitsbereich sind die geltenden Vorgaben angesichts der Sensibilität der verarbeiteten Daten verschärft. Insbesondere hier sollten sich Unternehmen mit den geltenden Vorschriften auseinandersetzen und diese im Interesse der betroffenen Personen, aber auch im eigenen Interesse befolgen.

### Literatur

- Abel RB (2018) Automatisierte Entscheidungen im Einzelfall. ZD 2018:304–307
- Deutscher Bundestag (2024) Gesetz zur Nutzung von Gesundheitsdaten zu gemeinwohlorientierten Forschungszwecken und zur datenbasierten Weiterentwicklung des Gesundheitswesens vom 22. März 2024 (Gesundheitsdatennutzungsgesetz – GDNG). BGBl 2024 I Nr 102
- Europäische Kommission (2020) Eine europäische Datenstrategie. COM 1 (66 final)
- Europäische Kommission (2025) Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act). Annex zu C (884 final)
- Europäischer Gerichtshof (2023) Urteil vom 07.12.2023, C-634/21, ECLI:EU:C:2023:957
- Europäisches Parlament, Rat der Europäischen Union (2016) Verordnung (EU) 2016/679 des Europäischen

- Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) ABl L, 119/1, 4. Mai 2016
- Europäisches Parlament, Rat der Europäischen Union (2024) Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) ABl L, 2024/1689, 12. Juli 2024
- Europäisches Parlament, Rat der Europäischen Union (2025) Verordnung (EU) 2025/327 des Europäischen Parlaments und des Rates vom 11. Februar 2025 über den europäischen Gesundheitsdatenraum sowie zur Änderung der Richtlinie 2011/24/EU und der Verordnung (EU) 2024/2847 ABl L, 2025/327, 5. März 2025
- Keber (2024) Verbot mit Erlaubnisvorbehalt nach Art. 6 Abs. 1 und Art. 9 Abs. 1 DS-GVO. In: Schwartmann R, Keber T, Zenner K (Hrsg.) KI-VO. Leitfaden für die Praxis, 2. Aufl. CF Müller, Heidelberg, S. 208–212
- Köhler M (2025) Das Verbot der automatisierten Einzelentscheidung und der hinreichende Einfluss des Menschen. EuDIR 01/2025:16–23
- Krämer W (2020) Die Rechtmäßigkeit der Nutzung von Scorewerten. NJW 8:497–502
- Paal B (2023) Art 22 DS-GVO: Kreditscoring vor dem EuGH. ZfDR 2:114–140
- Schwartmann R, Benedikt K, Köhler M (2024) Rechtspflege und demokratische Prozesse. In: Schwartmann R, Keber T, Zenner K (Hrsg.) KI-VO. Leitfaden für die Praxis, 2. Aufl. CF Müller, Heidelberg, S. 107–111
- Schwartmann R, Pottkämper E-M (2024) Hochrisiko-KI-Systeme gem. Art. 6 Abs. 1 KI-VO (Anhang I). In: Schwartmann R, Keber T, Zenner K (Hrsg.) KI-VO. Leitfaden für die Praxis, 2. Aufl. CF Müller, Heidelberg, S. 79–81
- Schwartmann R, Wasilewski D (2024) European Health Data Space (EHDS). In: Schwartmann R, Keber T, Zenner K (Hrsg.) KI-VO. Leitfaden für die Praxis, 2. Aufl. CF Müller, Heidelberg, S. 233–236
- Schwartmann R, Zenner K (2025) GPAI-Anwendungen auf dem Prüfstand: Die Regulierung der KI-VO entlang der Wertschöpfungskette. EuDIR 2025:3–9
- Schwartmann R, Benedikt K, Köhler M (2025) Expertenwissen KI und Datenschutz. Datakontext, Frechen
- Schwartmann R, Köhler M (2025) Rechtskonformer Einsatz von KI-Sprachsystemen im Unternehmen. In: Schwartmann R, Benedikt K, Reif Y (Hrsg.) Datenschutz im Internet, C.H.Beck, München, S. 711–729
- Stiftung Datenschutz (2023) Anonymisierung und Pseudonymisierung von Daten. CF Müller, Heidelberg
- Taeger J (2017) Verbot des Profiling nach Art 22 DS-GVO und die Regulierung des Scoring ab Mai 2018. RDV 1:3–9



# „KI als Medizin“ für das Gesundheitswesen?

Ein perspektivischer Artikel zum Einfluss von Künstlicher Intelligenz auf Gesundheit und Arbeit von medizinischem Personal

*Martin Roesler*

## Inhaltsverzeichnis

- 4.1 Handlungsbedarf und veränderte Rahmenbedingungen für die Medizin – 41**
- 4.2 Medizin ohne medizinisches Personal? – 42**
- 4.3 Entlastung von Routine- und Dokumentationsaufgaben – 43**
- 4.4 Konzentration auf Kernkompetenzen – mehr Zeit für Patienten und lebenslanges Lernen – 44**
- 4.5 Höhere Sicherheit bei der Entscheidungsfindung – psychische Entlastung für Ärztinnen und Ärzte – 44**
- 4.6 Risiken für medizinisches Personal – Überwachung, Leistungsdruck und Statusverlust – 45**
- 4.7 Gesundheitsberufe mutmaßlich unterschiedlich von Automatisierung betroffen – 46**
- 4.8 Die allgemeine Transformation des Arbeitsmarktes – 47**

**4.9 Fazit – Chancen und Fallstricke  
für Gesundheitsfachkräfte – 47**

**Literatur – 49**

## ■■ Zusammenfassung

Dieser Beitrag analysiert die potenziellen Auswirkungen generativer KI – insbesondere Large-Language-Models (LLMs) – auf die Arbeitsbedingungen des medizinischen Personals in Deutschland. Ausgangspunkt ist eine Evidenzlage, die wachsendem Fachkräftemangel, administrativer Überlastung und demographisch bedingter Mehrarbeit eine signifikante Gefährdung der Versorgungsqualität attestiert. Vor diesem Hintergrund werden aktuelle Studien zu LLM-gestützten Schreib- und Kommunikationsassistenzsystemen ausgewertet, die eine messbare Reduktion der Dokumentationszeit und eine Verbesserung der Patientenedukation belegen. Als technologische Fallstudien dienen die Systeme Forward Care-Pod, H4D Consult Station und H3 Health Cube; sie repräsentieren unterschiedliche Automatisierungsgrade einer „Praxis-in-der-Box“ und illustrieren sowohl die funktionellen Potenziale als auch die Implementierungsrisiken KI-basierter Selbstbedienungskioske. Im Vergleich weist die Pflege- und Physiotherapie aufgrund ihres hohen Anteils manuell-körperorientierter Tätigkeiten aktuell eine geringere Substituierbarkeit auf, was systematische Übersichtsarbeiten zur Robotik- und KI-Adaption bestätigen. Abschließend wird argumentiert, dass KI ihr entlastendes und qualitätssteigerndes Potenzial nur entfalten kann, wenn technische Zuverlässigkeit, institutionelle Governance, qualifikationsadäquate Weiterbildung und eine gerechte Reinvestition der Effizienzgewinne in Patientenversorgung und berufliche Entwicklung gewährleistet sind.

## 4.1 Handlungsbedarf und veränderte Rahmenbedingungen für die Medizin

Die Arbeitsbedingungen für medizinisches Personal in Deutschland haben sich in den letzten Jahren erheblich verschlechtert. Die Kombination aus Fachkräftemangel, wachsender Bürokratie und steigender Komplexität me-

dizinischer Behandlungen führt zu einer immer höheren Arbeitsbelastung (Freese 2025). In deutschen Krankenhäusern ist der Zeitdruck massiv angestiegen, sodass Ärztinnen und Ärzte sowie Pflegekräfte oft weit über die reguläre Arbeitszeit hinaus tätig sind. Eine der größten Herausforderungen stellt der administrative Aufwand dar, der inzwischen einen erheblichen Teil der Arbeitszeit beansprucht. Die Dokumentationspflichten, Abrechnungsverfahren und regulatorischen Vorgaben binden wertvolle Ressourcen, die für die direkte Patientenversorgung fehlen.

Dies führt dazu, dass Personalengpässe durch Mehrarbeit kompensiert werden müssen, was eine zunehmende physische und psychische Belastung für die Beschäftigten bedeutet. Laut aktuellen Untersuchungen leidet bspw. ein hoher Anteil des medizinischen Personals unter chronischem Stress und Erschöpfung, was nicht nur deren Gesundheit gefährdet, sondern langfristig auch die Versorgungsqualität der Patientinnen und Patienten beeinträchtigen kann (Freese 2025).

Angesichts dieser Herausforderungen erscheint der Ruf nach technologischer Unterstützung immer plausibler. Künstliche Intelligenz (KI), insbesondere Large Language Models (LLMs) wie GPT-4, zeigen das Potenzial, medizinisches Personal von zeitraubenden Routineaufgaben zu entlasten. Diese Systeme können administrative Prozesse automatisieren, medizinische Dokumentationen effizienter gestalten und Ärztinnen und Ärzte mit schnellen, datenbasierten Analysen unterstützen. Initiale wissenschaftliche Arbeiten deuten darauf hin, dass eine erfolgreiche Integration von KI in den klinischen Alltag das Gesundheitswesen nachhaltig verändern könnte – bspw. durch präzisere Diagnosen (Liu et al. 2019), eine gezieltere Steuerung der Patientenversorgung (Chen et al. 2024) und eine deutliche Reduzierung der administrativen Tätigkeiten (Hazan et al. 2024). Damit könnte KI eine Antwort auf einen wichtigen Anteil der aktuellen strukturellen Probleme des Gesundheitssystems liefern und dazu beitragen, eine

qualitativ hochwertige Versorgung auch in Zukunft sicherzustellen.

## 4.2 Medizin ohne medizinisches Personal?

Bereits 2023 führte das US-Start-up Forward Health die sogenannten CarePods ein – hochmoderne, KI-gestützte Selbstbedienungskioske, die Routineuntersuchungen ohne direktes ärztliches Zutun ermöglichen sollten (Price 2024). Die CarePods sollten flächendeckend in öffentlichen Bereichen wie Einkaufszentren und Fitnessstudios installiert werden und boten Funktionen wie biometrische Scans, Bluttests und diagnostische Sensoren an. Der erste CarePod wurde im Roseville Galleria Einkaufszentrum in Roseville, Kalifornien, installiert.

Die Einführung dieser Technologie führte zu einer grundlegenden Veränderung der Rolle des medizinischen Personals. Ärztinnen und Ärzte überwachten die von den CarePods erhobenen Befunde aus der Ferne und griffen nur bei Bedarf aktiv ein. Adrian Aoun, CEO von Forward, zog einen Vergleich zum Geldautomaten: Ärzte würden nicht mehr 100 % der Versorgung übernehmen, sondern sich nur noch um die letzten 5 % oder die besonders komplexen Fälle kümmern, während der Rest vollautomatisch ablaufe (Dashevsky 2023).

Trotz des innovativen Ansatzes wurde die Technologie im Jahr 2024 wieder vom Markt genommen. Hauptgründe waren technische Probleme wie häufige Fehlfunktionen bei automatisierten Blutabnahmen und Fälle, in denen Patientinnen und Patienten in den Pods stecken blieben. Zudem war das Interesse der Patienten geringer als erwartet und das Unternehmen kämpfte mit logistischen Herausforderungen bei der Implementierung der Pods. Diese Faktoren führten zu finanziellen Schwierigkeiten und letztlich zur Schließung des Unternehmens (Price 2024).

Die französische Telemedizinplattform **H4D** verfügt mit der *Consult Station* über eine vergleichbare Kabine, die in ihrer Vollversion bis zu 14 vernetzte Instrumente – darun-

ter Stethoskop, Otoskop, EKG-Ableitung, Blutdruck- und Pulsoximeter – integriert und Ärztinnen und Ärzten eine synchrone Video-konsultation bei gleichzeitig automatisierter Erfassung sämtlicher Vital- und Befunddaten ermöglicht (Falgarone et al. 2022). Nach der gerichtlichen Liquidation des Unternehmens am 26. September 2024, wodurch rund 150 installierte Kabinen außer Betrieb gesetzt wurden (Djabari 2024), übernahm im März 2025 eine Lyoner Plattform den gesamten Hardware-Bestand; seit Mai 2025 sind erst wenige Einheiten – beginnend mit einer Station in Menton – wieder betriebsfähig, während der vollständige Re-Roll-out bis 2026 geplant ist (MaQuestionMedicale 2025). Funktional ähnelt die Consult Station des US-amerikanischen **Forward CarePod**, da beide Systeme eine begehbar „Praxis-in-der-Box“ mit integrierter Sensorik, automatisierter Datenerhebung und nachgelagerter ärztlicher Validierung bieten. Die wesentlichen Unterschiede liegen in der Arbeitslogik und dem Reifegrad: Beim CarePod übernahm eine proprietäre KI den Großteil der Triage und Entscheidungsfindung, während klinisches Personal nur die Endfreigabe leistet, wohingegen H4D konsequent auf eine Echtzeit-Telekonsultation setzt. Die französische Consult Station wird als telemedizinisch betreute Selbstmess-Station in Gebieten eingesetzt, in denen keine ausreichende hausärztliche Versorgung besteht.

Mit ihrem **H3 Health Cube „Medical Cabin“** erprobte das kanadische Unternehmen **UniDoc Health** einen Prototyp unter anderem in der italienischen Gemeinde Aliano. Es handelt sich hierbei um eine begehbar, rollstuhlgerechte Telemedizin-Kabine, die wie der CarePod über ein Blutabnahme verfügt (Coxworth 2025; Galaxy Security 2025). Bei dieser Kabine sollen alle Messwerte durch eine KI vorbewertet werden und damit soll die Entscheidungsfindung für die Ärztinnen und Ärzte vorbereitet werden. Regulatorisch fällt das System in Europa damit doppelt unter hochregulierte Kategorien: als vernetztes Medizinprodukt nach **EU-MDR 2017/745** und als

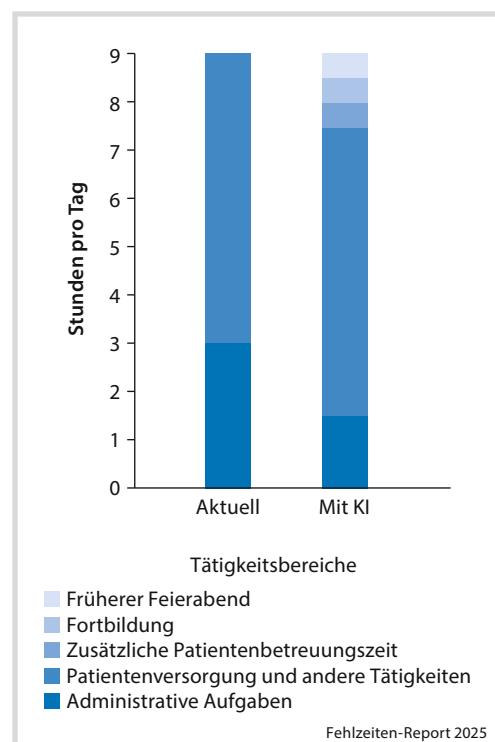
„High-Risk AI System“ im Sinne des EU-AI-Act (EU 2024/1689, Art. 6 Abs. 1 b), weil die KI ein sicherheitsrelevanter Bestandteil des Medizinprodukts ist. Daraus ergeben sich ergänzende Pflichten zu Risiko- und Datenmanagement, Bias-Kontrolle, menschlicher Kontrolle und technischer Dokumentation, die spätestens ab 2. August 2027 vollumfänglich nachgewiesen werden müssen (Angeloudi und Gillert 2025).

### 4.3 Entlastung von Routine- und Dokumentationsaufgaben

Ein zentrales Einsatzfeld von LLMs liegt in der **Automatisierung administrativer Routinetätigkeiten**. Insbesondere bei Dokumentationsaufgaben zeigen sich große Potenziale: Der Einsatz von LLMs als *Schreibassistent* für medizinisches Personal kann den für Dokumentation benötigten Zeitaufwand deutlich senken. Arztbriefe, Befundberichte oder Abrechnungen lassen sich mithilfe von KI deutlich schneller erstellen. Eine deutsche Studie untersuchte z.B. die Machbarkeit von Open-Source-LLMs zur klinischen Dokumentation und stellte fest, dass **93 % der vom LLM generierten Entwürfe für Arztberichte direkt verwendbar oder nur mit minimalen Korrekturen** waren – ein Hinweis auf das enorme Entlastungspotenzial solcher Systeme (Heilmeyer et al. 2024).

Weitere Beispiele aus der Forschung untermauern dieses Entlastungspotenzial. So berichteten Forscher in *The Lancet Digital Health* von einem LLM-Einsatz, der die Arbeitslast von Ärzten senkte und gleichzeitig die Patientenedukation verbessern konnte (Chen et al. 2023).

In einer aktuellen Untersuchung wurde GPT-4 genutzt, um Antworten auf Patientenanfragen im Klinikalltag zu formulieren. Die Ergebnisse zeigten, dass die KI-generierten Texte in **58 % der Fälle ohne Anpassung versandfertig** waren. Die beteiligten Ärzte berichteten von einer spürbaren Reduktion ihrer



■ Abb. 4.1 Eine mögliche Veränderung der täglichen Arbeitszeitverteilung von Arztinnen und Ärzten bei Einsatz von KI-Systemen. Die Grafik vergleicht die aktuelle Arbeitszeitstruktur (bei einem 9-Stunden-Arbeitstag) mit einem Szenario, in dem KI administrative Tätigkeiten halbiert. Die entstehende Zeiterparnis wird in diesem wünschenswerten Modell für zusätzliche Patientenbetreuung, Fortbildung und eine Verkürzung der täglichen Arbeitszeit genutzt. (Eigene Darstellung)

Arbeitsbelastung und weniger Burnout-Gefühlen (Chen et al. 2023).

Durch solche Automatisierungen gewinnen medizinische Fachkräfte wertvolle Zeit, die sie statt für Bürokratie wieder verstärkt für die Betreuung ihrer Patienten einsetzen können. Routine- und Dokumentationsaufgaben, die bislang als zeitraubende Pflicht galten, könnten so zumindest partiell automatisiert werden und den Weg für eine **patientenzentriertere Nutzung der ärztlichen Arbeitszeit** freimachen (■ Abb. 4.1).

## 4.4 Konzentration auf Kernkompetenzen – mehr Zeit für Patienten und lebenslanges Lernen

Werden Routineaufgaben von KI-Systemen übernommen, können sich Ärztinnen und Ärzte stärker auf ihre Kernkompetenzen fokussieren – die direkte medizinische Versorgung und die persönliche Arzt-Patienten-Interaktion. Der gewonnene Freiraum erlaubt **mehr persönliche Gespräche und Zuwendung zu den Patientinnen und Patienten**, was die Vertrauensbeziehung stärkt und die Behandlungsqualität erhöht. Studien deuten darauf hin, dass wenn Ärztinnen und Ärzte weniger Zeit mit Bürokratie verbringen müssen, ihre **Arbeitszufriedenheit** steigt (Gallani 2024). KI könnte somit helfen, die Medizin wieder „menschlicher“ zu machen, indem sie Zeit für die Aspekte schafft, die nur Menschen leisten können, wie bspw. körperliche Präsenz und visuell emphatische Kommunikation. Zudem ist davon auszugehen, dass die Reduktion der als illegitim wahrgenommenen bürokratischen Arbeitslast den Stress des medizinischen Personals bei der Arbeit senken kann (Semmer et al. 2013).

Zudem eröffnet die Entlastung von Routinearbeiten Raum für **lebenslanges Lernen**. Medizinisches Wissen verdoppelt sich bekanntlich in immer kürzeren Zyklen, und es ist für Praktizierende eine Herausforderung, mit den neuesten Entwicklungen Schritt zu halten (Kundu 2021). Wenn KI gewisse Aufgaben abnimmt, können Ärztinnen und Ärzte die frei gewordene Zeit nutzen, um sich fortzubilden und neue Forschungserkenntnisse in ihre Praxis zu integrieren. Die Anpassungsfähigkeit des medizinischen Personals steigt, weil kontinuierliche Weiterbildung besser mit dem Arbeitsalltag vereinbar wird. Langfristig dürfte dies dazu beitragen, dass die Berufsangehörigen kompetent bleiben im Umgang mit neuen Therapieoptionen, Leitlinien und Technologien – inklusive der KI selbst (Bundesärztekammer 2025). Insgesamt bedeutet die Kon-

zentration auf Kernkompetenzen also: **mehr Zeit für den Patienten und für die eigene Weiterbildung**, was sowohl der Versorgungsqualität als auch der beruflichen Entwicklung zugutekommt.

## 4.5 Höhere Sicherheit bei der Entscheidungsfindung – psychische Entlastung für Ärztinnen und Ärzte

Ein weiterer Vorteil des KI-Einsatzes liegt in der **Unterstützung klinischer Entscheidungen**. LLMs können als eine Art „zweite Meinung“ dienen und bspw. Ärztinnen und Ärzte bei komplexen diagnostischen oder therapeutischen Überlegungen unterstützen. Die Modelle sind in der Lage, enorme Wissensbestände in Sekunden zu durchsuchen – von Leitlinien über aktuelle Studien bis hin zu seltenen Fallberichten – und daraus fundierte Hinweise abzuleiten. So kann KI dazu beitragen, Fehler zu vermeiden, indem sie beispielsweise auf Wechselwirkungen von Medikamenten hinweist oder relevante Vorerkrankungen eines Patienten mitdenkt, die im hektischen Klinikalltag übersehen werden könnten. Eine **bessere Verfügbarkeit von Information** und evidenzbasiertem Wissen durch KI bedeutet, dass Entscheidungen breiter abgestützt und geprüft werden. Forschungsergebnisse unterstützen diesen Ansatz: In einer Übersichtsarbeit zeigte sich, dass fortgeschrittene KI-Systeme (etwa mit multimodaler Datenverarbeitung) evidenzbasierte Entscheidungen verbessern und sogar bei **seltenen Krankheiten präzise Diagnosen und Prognosen** ermöglichen konnten (Du et al. 2024).

Konkrete Beispiele belegen das Potenzial zur Erhöhung der Entscheidungssicherheit. Eine in *JAMA Network Open* veröffentlichte randomisierte Studie untersuchte bspw., wie sich ein LLM auf die diagnostische Leistung von Ärzten auswirkt. Dabei erzielte das KI-Modell (GPT-4) – für sich allein genommen – **eine signifikant höhere diagnostische Treffsicherheit** als die Ärzte, die konventionelle

Hilfsmittel nutzten (Goh et al. 2024). Mit anderen Worten: Die KI identifizierte korrekte Diagnosen in Fällen, die menschliche Kolleginnen und Kollegen übersahen. Interessanterweise verbesserten sich jedoch die Diagnosen der Ärzte **nicht automatisch**, nur weil ihnen das KI-Tool zur Verfügung stand. Das LLM hatte unter diesen Studienbedingungen zwar das Potenzial, bessere Diagnosen zu stellen als die Ärztinnen und Ärzte, wobei sich aber allein durch die Nutzungsmöglichkeit der KI dieses Potenzial zur besseren Diagnosefindung nicht realisierte. Zudem ist anzumerken, dass es sich hierbei um eine einzelne Studie handelt, in der viele ärztliche Arbeitsschritte bereits vorwegengenommen waren, um dem LLM seine Arbeit zu ermöglichen.

Dies unterstreicht jedoch, dass die optimale Zusammenarbeit zwischen Menschen und KI erst noch entwickelt werden muss. Dennoch verdeutlicht das Ergebnis das enorme Potenzial von KI, die klinische Entscheidungsfindung in Zukunft zu unterstützen. *KI-gestützte Diagnosesysteme* könnten als Sicherheitsnetz dienen, das ärztliche Entscheidungen auf Plausibilität prüft und bei Unsicherheiten zusätzliche Vorschläge für Differentialdiagnosen oder Behandlungsverläufe liefert. Richtig eingesetzt, haben LLMs damit das Potenzial, die **Patientensicherheit zu erhöhen** und Ärzte vor folgenschweren Irrtümern zu bewahren, indem sie Wissen jederzeit verfügbar machen und den Blick über den Tellerrand erleichtern.

#### 4.6 Risiken für medizinisches Personal – Überwachung, Leistungsdruck und Statusverlust

Den vielfältigen Chancen stehen auch **ernsthafte Risiken** gegenüber. Ein großes Bedenken ist die mögliche **Überwachung des medizinischen Personals** durch KI-Systeme. Moderne digitale Werkzeuge ermöglichen Arbeitgebern, jeden Arbeitsschritt und jede Entscheidung von Ärztinnen und Ärzten detailliert nachzuverfolgen (Ferguson 2025). So könnte

z. B. aufgezeichnet werden, wie lange ein Arzt für bestimmte Patientenakten braucht, welche Empfehlungen der KI er übernimmt oder verwirft etc. Dieses *gläserne Arztzimmer* schürt die Sorge, dass KI primär als **Kontrollinstrument** dient anstatt als Unterstützung. Wenn Algorithmen zur Leistungsbewertung herangezogen werden, könnte dies zu einem Klima ständiger Beobachtung führen, in dem sich Ärzte weniger frei fühlen, ihren klinischen Ermessensspielraum auszuschöpfen.

Zudem besteht die Gefahr, dass Effizienzgewinne durch KI in erhöhten **Leistungsdruck** münden. Wenn dokumentationsintensive Aufgaben schneller erledigt werden, könnten Klinikleitungen erwarten, dass Ärzte im Gegenzug mehr Patienten in kürzerer Zeit behandeln. Tatsächlich weisen Gesundheitsexperten darauf hin, dass frei werdende Zeit im aktuellen System leicht von den Arbeitgebern vereinnahmt wird – sprich: Es werden einfach mehr Fälle pro Tag verlangt (Gallani 2024). KI könnte so **das bereits überfordernde Arbeitstempo weiter steigern**, anstatt es zu begrenzen. Statt Entlastung zu spüren droht dem Personal eine weitere Verdichtung des Arbeitstages. In Kombination mit einer engmaschigen Überwachung würde dies die mentale Belastung zusätzlich erhöhen.

Auch die **ärztliche Autonomie** könnte unter verstärktem KI-Einsatz leiden. Wenn algorithmische Systeme Behandlungspfade vorgeben oder Versicherungen KI-Entscheidungshilfen nutzen, um Therapieentscheidungen zu beeinflussen, schrumpft der unabhängige Entscheidungsspielraum der Ärzte. Die Gefahr besteht, dass sich die Rolle des Arztes vom freien Heilkundigen zum Ausführenden von KI-Empfehlungen wandelt. Damit verbunden ist schließlich die Sorge vor einem **Statusverlust des Arztberufs**. Sollte die KI in immer mehr Kernbereichen ebenso kompetent – oder in manchen Aspekten sogar kompetenter – erscheinen als menschliche Ärzte, könnte dies langfristig das Ansehen und Selbstverständnis der Profession untergraben. Einige Beobachter warnen, dass eine übermäßige KI-gesteuerte Standardisierung und Kontrolle die **Arbeits-**

**zufriedenheit und die berufliche Identität** von Ärzten beeinträchtigen könnte (Ferguson 2025). Der Arzt würde Gefahr laufen, vom souveränen Entscheider zum überwachten „Systemnutzer“ degradiert zu werden. Diese Perspektive der Entmenschlichung und Entwertung macht deutlich, dass der KI-Einsatz mit Bedacht gesteuert werden muss, um die **ärztliche Autonomie und Würde** des Berufsbildes zu erhalten.

Ein weniger besprochenes Problem ist die **Fähigkeitsatrophie** (engl.: *skill atrophy*), bei der die ständige Nutzung von KI-Assistenzsystemen dazu führen kann, dass medizinisches Personal seine eigenen diagnostischen und analytischen Fähigkeiten vernachlässigt und somit abbaut. So weisen Studien bspw. darauf hin, dass die Abhängigkeit von KI-Systemen das eigene Urteilsvermögen beeinträchtigen kann (Macnamara et al. 2024).

Rechtlich gesehen bleibt die Verantwortung für medizinische Entscheidungen beim Fachpersonal. Wenn Ärztinnen und Ärzte gut formulierte Fehler in von KI generierten Inhalten übersehen, können sie haftbar gemacht werden. Der Deutsche Ethikrat betont, dass KI-gestützte Entscheidungssysteme so gestaltet sein sollten, dass ihre Ergebnisse transparent und nachvollziehbar sind, um Fehlentscheidungen zu minimieren (Deutscher Ethikrat 2023). Aktuell sind KI-Prozesse aufgrund der ihnen innenwohnenden Black-Box-Problematik nicht transparent.

Es ist daher unerlässlich, dass auch Gesundheitspersonal trotz technologischer Unterstützung seine Fachkompetenz kontinuierlich pflegt und KI-Ergebnisse kritisch zu hinterfragen lernt, um sowohl die Patientensicherheit als auch die eigene rechtliche Absicherung zu gewährleisten.

## 4.7 Gesundheitsberufe mutmaßlich unterschiedlich von Automatisierung betroffen

Vor der Ära der generativen KI wurde die skalierbare Substitution von ärztlicher Arbeit als gering eingeschätzt (Davenport und Kalakota 2019). Ironischerweise steht aktuell das logische Denken durch Modelle mit *advanced reasoning* (deutsch: fortgeschrittenes Schlussfolgern), das lange als höchstes Gut und Alleinstellungsmerkmal der Menschen galt, vor der Automation (Bastian 2024). Gegenüber der ärztlichen Tätigkeit ist derzeit davon auszugehen, dass zum Beispiel Pflegekräfte und die Arbeit in der Physiotherapie in den kommenden Jahren in geringerem Maße von generativer KI betroffen sein werden. Dies ist vor allem im hohen Anteil praktisch-körpernaher Tätigkeiten begründet, die sich bislang kaum automatisieren oder sprachbasiert unterstützen lassen. So betonen systematische Übersichtsarbeiten, dass Aufgaben wie körperliche Betreuung, manuelle Mobilisation, Bewegungskorrektur oder die situativ-emotionale Interaktion mit Patienten wesentlich bleiben und durch KI nur in sehr begrenztem Umfang ergänzt werden können (Adeyemo et al. 2025; Vaniya et al. 2024). Zwar gibt es Anwendungen, etwa zur automatisierten Dokumentation oder zur Unterstützung bei der Trainingsplanung, doch bleibt der Charakter der Arbeit überwiegend menschlich-praktisch geprägt. Demgegenüber sind ärztliche Tätigkeiten – insbesondere in Bereichen wie Diagnostik, Dokumentation und Entscheidungsunterstützung – in höherem Maß digitalisierbar und damit empfänglicher für generative KI-Systeme. Es ist daher anzunehmen, dass es eher zu einem unterstützenden, punktuellen KI-Einsatz in Pflege und Physiotherapie kommt, während sich für die ärztliche Rolle mittelfristig stärkere strukturelle Veränderungen abzeichnen. Für chirurgische Tätigkeiten lässt sich dies ähnlich begründen. Wenn die Robotik sich ebenso rasant entwickelt wie die

KI, verliert diese Argumentation natürlich wesentliche Teile ihrer Substanz.

## 4.8 Die allgemeine Transformation des Arbeitsmarktes

Die Auswirkungen von KI beschränken sich nicht auf den Gesundheitsbereich, sondern sind Teil einer umfassenden **Transformation des gesamten Arbeitsmarktes**. Der *Future of Jobs Report 2025* des Weltwirtschaftsforums (WEF) prognostiziert, dass bis 2030 etwa **22 % aller Arbeitsplätze** weltweit grundlegende Veränderungen durch Automatisierung und neue Technologien erfahren werden (WEF 2025). Gleichzeitig sollen im selben Zeitraum rund 170 Mio. neue Jobs entstehen (ca. 14 % der heutigen Beschäftigung), während etwa 92 Mio. Stellen wegfallen könnten (8 % der heutigen Jobs). Trotz eines erwarteten Nettozuwachses von Arbeitsplätzen bedeuten diese Verschiebungen erhebliche Umbrüche für viele Berufsfelder (Figure 2.1, WEF 2025).

Ein Blick in andere Branchen zeigt, dass KI und Automation dort bereits fest verankert sind und teils **tiefgreifende Veränderungen** bewirkt haben. So gehören einfache Bürotätigkeiten und Servicejobs – etwa **Kassierer, Datenerfasser oder administrative Sachbearbeiter** – zu den am stärksten schrumpfenden Berufsgruppen durch den Einsatz von KI (Figure 2.2, WEF 2025). Routineaufgaben in Banken, Versicherungen oder im Handel werden zunehmend von Algorithmen oder Robotern übernommen (man denke an Chatbots im Kundendienst oder Automatisierung in der Buchhaltung). Demgegenüber wachsen technologiegetriebene Berufe rasant: Spezialisten für *Big Data*, KI-Entwicklung oder Automatisierungstechnik sind global stark nachgefragt. Diese Verschiebung – **Wegfall repetitiver Tätigkeiten, Entstehung neuer hochqualifizierter Jobs** – zeichnet sich branchenübergreifend ab.

Im Gesundheitswesen rechnet das WEF aufgrund der Alterung der Gesellschaft mit

einer steigenden Nachfrage nach Gesundheitsdienstleistern – Pflegekräfte und medizinische Fachkräfte werden weiterhin gebraucht und ihre Anzahl soll sogar überdurchschnittlich wachsen (Figure 2.3 WEF 2025). Szenarisch könnte dies bedeuten, dass sich Berufsprofile im medizinischen Bereich verändern: Administrative Routinearbeiten werden reduziert, während neue Aufgaben rund um die **digitale Gesundheitsversorgung** entstehen. Möglicherweise werden zukünftig vermehrt Spezialisten benötigt, die KI-Systeme warten, medizinische Daten kuratieren oder die Schnittstelle zwischen Technik und klinischer Praxis managen. Ärztinnen, Ärzte und Pflegekräfte werden sich mutmaßlich stärker auf Tätigkeiten konzentrieren, die manuelle Fähigkeiten und menschliche Interaktion erfordern. Insgesamt ähnelt das Spannungsfeld im Gesundheitsbereich dem anderer Branchen: KI wird Jobs in der Gesundheit vermutlich **verändern, aber nicht verschwinden lassen**. Es kommt zu einer Neuverteilung von Aufgaben, bei der datenbasierte Tätigkeiten automatisiert und menschliche Fähigkeiten in anderen Bereichen umso wichtiger werden.

## 4.9 Fazit – Chancen und Fallstricke für Gesundheitsfachkräfte

Die Einführung von KI und LLM-Technologien im Gesundheitswesen wird die Arbeitswelt transformieren.

Auf der Habenseite stehen potenziell eine **Entlastung von bürokratischen Lasten**, mehr Zeit für Patientengespräche, präzisere Entscheidungen durch Unterstützungssysteme und letztlich auch die Möglichkeit, den Beruf wieder näher an seine kernberuflichen Ideale zu rücken. Erste Studien zeigen, dass KI-Assistenz die Dokumentationsqualität steigern und die Arbeitszufriedenheit erhöhen kann, wenn sie richtig eingesetzt wird. LLMs als digitale Helfer könnten das Arbeiten in der Medizin effizienter gestalten und dazu beitragen, Fehler zu reduzieren. Diese Technologien sollten allerdings als *Erweiterung* der mensch-

lichen Fähigkeiten verstanden werden – sie können insbesondere administrative Routinearbeit abnehmen und Informationen liefern, benötigen derzeit aber noch bei Fehleranfälligkeit besonders geschulte menschliche Supervision.

Auf der Sollseite stehen Risiken wie **Überautomatisierung, Überwachung und Entmenschlichung**. Ohne Leitplanken könnte KI den ökonomischen Druck auf Ärztinnen und Ärzte weiter erhöhen oder ihre Autonomie beschneiden. Die ärztliche Kunst droht beschädigt zu werden, wenn sich Mediziner unkritisch auf algorithmische Vorgaben verlassen oder unter dem Diktat weiterer Effizienzsteigerungen stehen.

Um KI gewinnbringend einzusetzen, bedarf es einer **aktiven Gestaltung** des Wandels. Die private Schattennutzung von LLM durch Mitarbeiterinnen und Mitarbeiter über inoffizielle Kanäle führt zu massiven datenschutzrechtlichen Problemen (Netskope 2025). Die Politik und Gesundheitsorganisationen sind gefordert, alles Notwendige zu veranlassen – von pragmatischem Datenschutz über Haftungsfragen bis hin zu Qualitätsstandards –, damit KI-Anwendungen schnell, zuverlässig und sicher in den Klinikalltag integriert werden können. Arbeitsmodelle sollten so angepasst werden, dass Effizienzgewinne durch KI nicht zu Lasten der Beschäftigten gehen, sondern zu deren tatsächlichen Entlastung und einer besseren Patientenbetreuung genutzt werden (Gallani 2024). Gleichzeitig muss in die **Aus- und Weiterbildung** investiert werden. Ärzte und Pflegende sollten im Umgang mit den neuen Technologien geschult werden, um deren Möglichkeiten kritisch und kompetent nutzen zu können. Die Vermittlung von KI-Kompetenzen und digitaler Gesundheitskompetenz sollte fest in der medizinischen Ausbildung verankert werden (Akinrinmade et al. 2023), damit das Personal auf die veränderten Anforderungen vorbereitet ist. Generative KI sollte trotz umfangreicher Risiken, Halluzinationen und Datenschutz (WHO 2025) zeitnah flächendeckend als *unterstützender Partner im Gesundheitswesen* eingeführt werden – als ein

Werkzeug, das dem medizinischen Personal bereits jetzt lästige Routinearbeit abnimmt und wertvolle Informationen liefert, während die finale Verantwortung und die menschliche Zuwendung beim ärztlichen und beim pflegerischen Personal verbleiben. Diese Redundanz schafft Sicherheit und ermöglicht sowohl die Technologie im Betrieb weiterzuentwickeln als auch bereits nutzbare Strukturen für sich abzeichnende zukünftige Entwicklungssprünge zu etablieren. Die Alternative ist die ineffizientere, unsichere und unkontrollierte private Nutzung.

Damit sich bei der Einführung der KI nicht die gleichen Fehler wiederholen wie bei der Digitalisierung des Gesundheitswesens, bedarf es einer veränderten Herangehensweise. Es wäre wünschenswert, wenn sich gesellschaftlich ein Mentalitätswechsel vollziehen würde. Statt eines defensiven Primats der Wahrung abstrakter Rechte sollte sich bspw. eher eine Art „mutiges Ingenieursdenken“ etablieren. Technische Möglichkeiten, praktischer Nutzen und konkretes Schadensrisiko sollten bei der Umsetzung im Vordergrund stehen. Gesundheitsbetriebe sind hierbei in der Pflicht, jetzt flächendeckend ihre Mitarbeiterinnen und Mitarbeiter in der Nutzung von generativer KI zu schulen, um mehr Rechtssicherheit zu erlangen und das Potenzial für gesteigerte Arbeitsqualität und -effektivität zu nutzen. So sollte man bspw. an einem Computerarbeitsplatz mittlerweile selbstverständlich Zugriff auf ein aktuelles großes Sprachmodell haben und die dort arbeitende Person die hinreichenden kritischen Fähigkeiten, um die gelieferten Information angemessen hinterfragen zu können.

**Disclaimer:** Dieser Artikel wurde unter Einbeziehung von ChatGPT Modell o3, einem KI-Sprachmodell von OpenAI, erstellt. Obwohl ChatGPT bei der Strukturierung, Recherche und Verfassung von Inhalten unterstützt hat, wurden alle Informationen durch den Autor sorgfältig überprüft und gegebenenfalls angepasst, um ihre Richtigkeit und Relevanz sicherzustellen.

## Literatur

- Adeyemo A, Coffey A, Kingston L (2025) Utilisation of robots in nursing practice: an umbrella review. *BMC Nurs* 24(1):247. <https://doi.org/10.1186/s12912-025-02842-2>
- Akinrinmade AO, Adebile TM, Ezuma-Ebong C, Bolaji K, Ajupo A, Adigun AO, Mohammad M, Dike JC, Okobi OE (2023) Artificial Intelligence in Healthcare: Perception and Reality. *Cureus* 15(9):e45594. <https://doi.org/10.7759/cureus.45594>
- Angeloudi JT, Gillert JE (2025) The EU AI act and medical devices: innovation or insanity ahead? *Healthcare & life sciences Blog*. <https://healthcarelifesciences.bakermckenzie.com/2025/05/19/the-eu-ai-act-and-medical-devices-innovation-or-insanity-ahead/>. Zugegriffen: 19. Mai 2025
- Bastian M (2024) OpenAI unveils five-level AI scale, aims to reach level 2 soon. *The Decoder*. <https://the-decoder.com/openai-unveils-five-level-ai-scale-aims-to-reach-level-2-soon/>. Zugegriffen: 12. Juli 2024
- Bundesärztekammer (2025) Stellungnahme „Künstliche Intelligenz in der Medizin“. [https://www.bundesaerztekammer.de/fileadmin/user\\_upload/wissenschaftlicher-beirat/Veroeffentlichungen/KI\\_in\\_der\\_Medizin\\_SN\\_neu.pdf](https://www.bundesaerztekammer.de/fileadmin/user_upload/wissenschaftlicher-beirat/Veroeffentlichungen/KI_in_der_Medizin_SN_neu.pdf). Zugegriffen: 21. Febr. 2025
- Chen S, Guevara M, Moningi S, Hoebers F, Elhalawani H, Kann BH, Chipidza FE, Leeman J, Aerts HJWL, Miller T, Savova GK, Mak RH, Lustberg M, Afshar M, Bitterman DS (2023) The impact of responding to patient messages with large language model assistance (No arXiv:2310.17703). arXiv. <https://doi.org/10.48550/arXiv.2310.17703>
- Chen S, Guevara M, Moningi S, Hoebers F, Elhalawani H, Kann BH, Chipidza FE, Leeman J, Aerts HJWL, Miller T, Savova GK, Gallifant J, Celi LA, Mak RH, Lustberg M, Afshar M, Bitterman DS (2024) The effect of using a large language model to respond to patient messages. *Lancet Digit Health* 6(6):e379–e381. [https://doi.org/10.1016/S2589-7500\(24\)00060-8](https://doi.org/10.1016/S2589-7500(24)00060-8)
- Coxworth B (2025) H3 Health Cube „self-contained connected clinic“ ships to first clients. *New Atlas*. <https://newatlas.com/imaging-diagnostics/h3-health-cube-ships-first-units/>. Zugegriffen: 7. Mai 2025
- Dashevsky J (2023) Examining forward health's bold leap into automated primary care. *Healthcare huddle*. <https://www.healthcarehuddle.com/p/examining-forward-healths-bold-leap-automated-primary-care>. Zugegriffen: 6. Juni 2025
- Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. *Future Health* 6(2):94–98. <https://doi.org/10.7861/futurehosp.6-2-94>
- Deutscher Ethikrat (2023) Stellungnahme: Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Deutscher Ethikrat. <https://www.ethikrat.org/publikationen/stellungnahmen/mensch-und-maschine/>. Zugegriffen: 6. Juni 2025
- Djabari M (2024) H4D, pionnière des cabines de téléconsultation, en liquidation judiciaire. <https://www rtl.fr/actu/debats-societe/h4d-pionniere-des-cabines-de-teleconsultation-en-liquidation-judiciaire-7900426772>. Zugegriffen: 6. Juni 2025
- Du X, Zhou Z, Wang Y, Chuang Y-W, Yang R, Zhang W, Wang X, Zhang R, Hong P, Bates DW, Zhou L (2024) Generative large language models in electronic health records for patient care since 2023: a systematic review. *medRxiv*, 2024.08.11.24311828. <https://doi.org/10.1101/2024.08.11.24311828>
- Falgarone G, Bousquet G, Wilmet A, Brizio A, Faure V, Guilloet C, Baudino F, Roque I, Mayol S, Pamoukdjian F (2022) A Teleconsultation device, consult station, for remote primary care: Multisite prospective cohort study. *J Med Internet Res* 24(5):e33507. <https://doi.org/10.2196/33507>
- Ferguson M (2025) AI's invasion: Will it diminish physician autonomy even further? AI's invasion: Will it diminish physician autonomy even further? <https://opentools.ai/news/ais-invasion-will-it-diminish-physician-autonomy-even-further>. Zugegriffen: 18. Jan. 2025
- Freese H-J (2025) MB-Monitor 2024: Hohe Belastung, unzureichende Personalausstattung und Zunahme an Gewalterfahrungen | Marburger Bund Bundesverband. <http://www.marburger-bund.de/bundesverband/themen/marburger-bund-umfragen/mb-monitor-2024-hohe-belastung-unzureichende>. Zugegriffen: 6. Febr. 2025
- Galaxy Security (2025) Unidoc H3 health cube—Unicheck virtual clinics. [https://www.galaxysecurity.com/en/product/7010?utm\\_source=chatgpt.com](https://www.galaxysecurity.com/en/product/7010?utm_source=chatgpt.com). Zugegriffen: 6. Juni 2025
- Gallani S (2024) Can AI save physicians from burnout? | working knowledge. Harvard Business School. <https://www.library.hbs.edu/working-knowledge/can-ai-save-physicians-from-burnout>. Zugegriffen: 6. Juni 2025
- Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, Cool JA, Kanjee Z, Parsons AS, Ahuja N, Horvitz E, Yang D, Milstein A, Olson APJ, Rodman A, Chen JH (2024) Large language model influence on diagnostic reasoning: a randomized clinical trial. *Jama Netw Open* 7(10):e2440969. <https://doi.org/10.1001/jamanetworkopen.2024.40969>
- Hazan E, Madgavkar A, Chui M, Smit S, Maor D, Dandona GS, Huyghues-Despointes R (2024) A new future of work: The race to deploy AI and raise skills in Europe and beyond
- Heilmeyer F, Böhringer D, Reinhard T, Arens S, Lysenko L, Haverkamp C (2024) Viability of open large language models for clinical documentation in German health care: real-world model evaluation stu-

- dy. JMIR Med Inform 12:e59617. <https://doi.org/10.2196/59617>
- Kundu S (2021) How will artificial intelligence change medical training? Commun Med 1(1):1–3. <https://doi.org/10.1038/s43856-021-00003-5>
- Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, Mahendiran T, Moraes G, Shamdas M, Kern C, Ledsam JR, Schmid MK, Balaskas K, Topol EJ, Bachmann LM, Keane PA, Denniston AK (2019) A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 1(6):e271–e297. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)
- Macnamara BN, Berber I, Çavuşoğlu MC, Krupinski EA, Nallapareddy N, Nelson NE, Smith PJ, Wilson-Delfosse AL, Ray S (2024) Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? Cogn Res Princ Implic 9:46. <https://doi.org/10.1186/s41235-024-00572-8>
- MaQuestionMedicale (2025) MaQuestionMedicale reprend le parc des 150 cabines de télémédecine H4D. MaQuestionMedicale téléconsultation médecin, un médecin à proximité. <https://www.maquestionmedicale.fr/maquestionmedicale-reprend-le-parc-des-150-cabines-de-telemecdecine-h4d/>. Zugegriffen: 5. Apr. 2025
- Netskope (2025) Netskope Threat Labs Bericht: Gesundheitswesen 2025. Netskope. <https://www.netskope.com/de/netskope-threat-labs/threat-labs-report-healthcare-2025>. Zugegriffen: 6. Juni 2025
- Price RTR (2024) Inside Forward's failed attempt to revolutionize the doctor's office with AI. Business Insider. <https://www.businessinsider.com/healthcare-startup-forward-shutdown-carepod-adrian-aoun-2024-11>. Zugegriffen: 10. Juni 2025
- Semmer NK, Jacobshagen N, Meier LL, Elfering A, Kälin W, Tschan F (2013) Psychische Beanspruchung durch illegitime Aufgaben. In: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Junghanns G, Morschhäuser M (Hrsg) Immer schneller, immer mehr: Psychische Belastung bei Wissens- und Dienstleistungsarbeit. Springer, S 97–112 [https://doi.org/10.1007/978-3-658-01445-2\\_5](https://doi.org/10.1007/978-3-658-01445-2_5)
- Vaniya J, Gandhi NV, Patel KP, Bhatt C, Kaur KP (2024) Role and scope of artificial intelligence in physiotherapy: a literature review. Res J Med Sci 18(12):399–412. <https://doi.org/10.36478/makrjms.2024.12.399.412>
- WEF – World Economic Forum (2025) The future of jobs report 2025. <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>. Zugegriffen: 10. Juni 2025
- WHO – World Health Organisation (2025) Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. <https://www.who.int/publications/item/9789240084759>. Zugegriffen: 10. Juni 2025

# Von der Personalplanung bis zum Gesundheitsmanagement: Nutzungsmöglichkeiten von künstlicher Intelligenz im Personalwesen

Alexander Dregger

## Inhaltsverzeichnis

- 5.1 Warum sich ein Einsatz von KI im HR lohnt – 52**
- 5.2 Der Personallebenszyklus und mögliche Einsatzgebiete von KI – 53**
  - 5.2.1 Personalplanung – 54
  - 5.2.2 Personalmarketing – 55
  - 5.2.3 Personalauswahl – 55
  - 5.2.4 Onboarding von Personal – 58
  - 5.2.5 Personalentwicklung – 59
  - 5.2.6 Personalmanagement und Personalbindung – 60
- 5.3 Fallstricke bei der Implementierung von KI im HR-Bereich und mögliche Lösungen – 61**
- 5.4 Fazit – 62**
- Literatur – 62**

## ■■ Zusammenfassung

Nach aktuellen Zahlen des Digitalverbands Bitkom e. V. (2024a) nutzen 20 % der deutschen Unternehmen Künstliche Intelligenz (KI). Damit liegt Deutschland zwar oberhalb des EU-Durchschnitts von 13,48 % (Eurostat 2025), aber auf globaler Ebene hinkt Deutschland im Vergleich zu Ländern wie z. B. China mit einer Implementierungsquote von 50 % bei Großunternehmen hinterher (IBM und Morning Consult 2023). Nicht nur zwischen den Nationen gibt es Unterschiede – auch innerhalb von Unternehmen nutzt nicht jede Abteilung KI im gleichen Maße. So setzen nur 12 % der Personalabteilungen KI ein, im Gegensatz zu z. B. 34 % der Marketing & Sales-Abteilungen (McKinsey 2024).

Ein grundsätzliches Problem, das auch bei Human Resources bzw. HR-Abteilungen besteht, ist die mangelnde Kenntnis von passenden KI-Anwendungsfällen (Dregger 2022). Deshalb fokussiert sich der folgende Beitrag hierauf: Anhand eines Personallebenszyklus, beginnend bei der Personalplanung und -marketing und endend bei der Bindung von Personal, sollen zentrale Einsatzmöglichkeiten von KI für den HR-Bereich beschrieben werden. Dabei werden auch gesundheitsbezogene Themen berücksichtigt. Abschließend sollen Fallstricke bei der Implementierung von KI im HR aufgezeigt und mögliche Lösungen hierfür skizziert werden.

### 5.1 Warum sich ein Einsatz von KI im HR lohnt

Künstliche Intelligenz kann wesentlich dazu beitragen, vier zentrale Herausforderungen zu bewältigen, mit denen HR-Abteilungen konfrontiert sind:

Der Begriff „**War for Talents**“ wurde bereits in den späten 1990ern geprägt und ist heutzutage sogar noch relevanter für Unternehmen geworden. Da bestimmte Berufsbilder zunehmend an Attraktivität verlieren, bewerben sich immer weniger geeignete Personen

in diesen Bereichen. Dies betrifft insbesondere Tätigkeiten mit monotonen Aufgaben, wie bspw. die Überwachung von Produktionsanlagen, oder mit schwierigen Arbeitsbedingungen, wie bspw. pflegerische Tätigkeiten. Daneben ist in bestimmten Branchen die Nachfrage enorm gestiegen: So sollen bis 2040 663.000 IT-Fachleute (Bitkom e. V. 2024b) und bis 2049 280.000 Pflegekräfte fehlen (Statistisches Bundesamt 2024). Die erhöhte Nachfrage nach IT-Fachleuten erklärt sich einerseits aus einer erhöhten Nutzung von IT-Lösungen in Unternehmen, andererseits durch den demographischen Wandel, der zu einem altersbedingten Verlust von IT-Fachkräften führt. Nicht nur im IT-Bereich, sondern auch generell droht Unternehmen ein Wissensverlust durch die Abwanderung im Rahmen des demographischen Wandels. KI kann durch verschiedene Anwendungen im Bereich Personalmarketing, Recruiting, Personalbindung, aber auch Wissensmanagement diesen Herausforderungen vorbeugen (Abb. 5.1).

Eine weitere Herausforderung, bei der KI helfen kann, ist die **Beschleunigung und Personalisierung** von Prozessen. Personalisierung meint, dass KI individuelle Präferenzen von Personen erkennen kann und z. B. die Gestaltung der Sprache in Ausschreibungen, Formularen oder Chatbots anpassen kann. Insbesondere bei gefragten Berufen oder bei einer hohen Menge an Konkurrenz um einen beschränkten Bewerberpool in Ballungsgebieten ist es notwendig, dass Unternehmen zeitnah Personalprozesse wie eine Einstellung erledigen. Darüber hinaus kann ein Unternehmen durch die Personalisierung von Ausschreibungen und Berücksichtigung individueller Bedürfnisse bei weiteren Personalprozessen die Personalbindung erhöhen. Sowohl bei der Personalisierung als auch bei der Beschleunigung von Prozessen kann KI einen entscheidenden Mehrwert erzeugen, in dem sie z. B. Schritte automatisiert.

Ein weiterer, nicht zu unterschätzender Aspekt ist, dass der Einsatz von KI die Wahrnehmung der **Unternehmenskultur** auf Be-



Abb. 5.1 Darstellung des Anteils der KI-Nutzung in verschiedenen Unternehmensbereichen basierend auf McKinsey (2024) (eigene Darstellung)

Schließlich kann KI dazu beitragen, dass menschliche **Fehler und Belastungen** reduziert werden. Beispielsweise führen Unternehmen regelmäßig komplexe Planungsprozesse wie z. B. die Organisation von Arbeitsplänen oder die Personalbedarfsplanung durch. Ein datenbasierter und mit KI-Methoden unterstützter Prozess kann genau bei solchen Themen Führungskräfte, aber auch die HR-Abteilungen entlasten und Fehlern vorbeugen. Dabei ist KI aber ähnlich wie z. B. Recruiter und Recruiterinnen bei der Personalauswahl (Kanning 2015) nicht komplett fehlerfrei (Kleinman 2024). Deshalb ist eine Qualifizierung auch des HR-Personals in Bezug auf KI nötig, damit Fehler, die durch den Einsatz von KI entstehen, rechtzeitig erkannt werden und ein optimales Fehlerniveau erreicht wird. Letzteres bedeutet hierbei, dass beim System z. B. durch unvorhergesehene Gestaltung von Bewerbungsunterlagen Fehler auftreten können. Solche Fehler müssen durch das Fachpersonal im HR-Bereich erkannt werden und das KI-System muss nachtrainiert werden, um vermeidbare Fehler zu korrigieren.

werberseite verändern kann: Nutzt ein Unternehmen KI, so zeigt es einerseits, dass es offen für digitale Innovationen ist. Andererseits kann ein sinnvoll gestalteter Einsatz von KI das Gefühl von Wertschätzung bei der aktuellen oder in der Rekrutierung befindlichen Belegschaft erhöhen. Dies kann z. B. dadurch passieren, dass Bewerbungen durch KI schneller verarbeitet werden oder Unternehmen schneller bemerken, dass bei ihrem Personal gesundheitliche Probleme bestehen. Was jedoch nicht passieren darf, ist, dass durch die Nutzung von KI beim Personal der Eindruck entsteht, „eine Nummer im System“ zu sein (Dahm und Dregger 2019). Dies würde das Gefühl von Wertschätzung auf Seiten des Personals einschränken, was wiederum die Akzeptanz der KI-Systeme selbst gefährden und letztlich die Motivation der Belegschaft bzw. von Bewerbenden reduzieren könnte.

## 5.2 Der Personallebenszyklus und mögliche Einsatzgebiete von KI

Um einen möglichst guten Überblick über die verschiedenen Einsatzgebiete von KI im HR-Management zu geben, werden diese anhand von sechs Schritten im Personallebenszyklus beschrieben: Zu Beginn wird die **Personalplanung** und das **Personalmarketing** beschrieben, um anschließend näher auf die Themen **Personalauswahl** und **Onboarding** einzugehen. Die letzten beiden Schritte stellen die Themen **Personalentwicklung** und **Personalmanagement** dar (Tab. 5.1).

■ Tab. 5.1 Übersicht über die Einsatzmöglichkeiten von KI in verschiedenen Bereichen des Personalwesens

Personalbereiche	Planung	Marketing	Personal-auswahl	Onboarding	Personal-entwicklung	Personal-management und Personal-bindung
KI-Einsatzmöglichenkeiten	<ul style="list-style-type: none"> <li>Datenbasierte strategische Personalplanung</li> <li>Kurzfristige Personalplanung mittels automatisierter Schicht- und Einsatzplanungen</li> </ul>	<ul style="list-style-type: none"> <li>Automatisierte Generierung von Stellenausschreibungen</li> <li>Social Recruiting (Active Sourcing)</li> </ul>	<ul style="list-style-type: none"> <li>CV-Parsing</li> <li>KI-basierte Vorauswahl</li> <li>KI-basierte Job-Interviews</li> <li>KI-basiertes Bewerbertraining</li> </ul>	(RAG-basierte) Chatbots und Voicebots	<ul style="list-style-type: none"> <li>KI-basiertes Wissensmanagement</li> <li>Personalisierter Weiterbildung</li> <li>KI-basiertes Gesundheitsmonitoring</li> </ul>	<ul style="list-style-type: none"> <li>People Analytics</li> <li>KI-basiertes Retention Management</li> <li>KI-basierte Zeugniserstellung</li> </ul>

## 5.2.1 Personalplanung

Personalplanung umfasst zwei Ebenen: Die langfristige, strategische Personalbedarfsplanung sowie die kurzfristige, wiederkehrende Arbeits- bzw. Personaleinsatzplanung z. B. im Rahmen von Schichtplänen, um Abteilungen oder Projekte zu organisieren (Spengler et al. 2025). Auf beiden Ebenen kann KI weiterhelfen.

Die Ermittlung des langfristigen **Personalbedarfs** kann insbesondere bei größeren Unternehmen mit einem wenig datengetriebenen Vorgehen eine Herausforderung darstellen: Laut einer Studie des Instituts der deutschen Wirtschaft führen 56 % aller Unternehmen keine strategische Personalplanung über mehrere Jahre hinweg durch (Stippler et al. 2019). Zumindest bei Großunternehmen mit mehr als 1.000 Beschäftigten führen immerhin 60 % eine jährliche Personalplanung durch (Kirchherr et al. 2024); hinsichtlich einer mehrjährigen, langfristigen Personalplanung sehen aber auch hier die Zahlen schlechter aus. Selbst wenn eine solche Planung durchgeführt wird, wird anstelle eines datengetriebenen Vorgehens mithilfe des „Bauchgefühls“ der Führungskräfte und Exceltabellen (Jäger und Wickel-Kirsch 2017) einmal im Jahr eine Abfrage des Perso-

nalbestands durchgeführt, um Zielgrößen für das gesamte Unternehmen festzulegen. Eine solches Vorgehen ist kritisch zu sehen, da z. B. Probleme bei dynamischen, kurzfristigen Änderungen gibt und hierdurch Personalengpässe entstehen können (Ahlers et al. 2020). Zudem sorgt der Mangel an systematischer Personalplanung für Unzufriedenheit in den Unternehmen (Jäger und Wickel-Kirsch 2017), da das übrige Personal die Engpässe kompensieren muss. KI kann genau hier ansetzen und basierend auf Auftragsdaten, Mitarbeiterdaten zu Qualifikationen, historischen Arbeitsplänen, Zahlen zu Krankenständen und vergangenen Personalplanungen eine Prognose erstellen, wie viel Personal in den kommenden Jahren voraussichtlich benötigt wird. In diesen Prozess können auch die Meinungen bzw. Schätzungen der Führungskräfte systematisch integriert werden – entweder bei der Generierung der Prognose selbst oder bei der Bewertung der Prognose. Solche Prognosen können dabei mit deutlich höherer Dynamik erstellt werden und zudem kann man auch Szenarien hiermit analysieren.

Eine weitere Ebene der Personalplanung ist die kurzfristige Personaleinsatzplanung mithilfe von Schicht- und Einsatzplänen des Personals. Das Prinzip ist ähnlich wie bei der

langfristigen Planung, jedoch können hier andere Datenquellen zum Einsatz kommen. Beispielsweise muss die Verfügbarkeit von Ressourcen wie z. B. Ersatzteilen in einer Werkstatt oder Operationssälen in Krankenhäusern berücksichtigt werden, um eine regelmäßige Vorhersage zu gewährleisten. Insbesondere eine KI-basierte kurzfristige Personalplanung kann die Führungskräfte deutlich entlasten, da solche Pläne mit hoher Regelmäßigkeit erstellt werden müssen. Zentral bei der Nutzung von KI ist, dass die notwendigen Daten in der benötigten Aktualität und Qualität digital vorliegen.

## 5.2.2 Personalmarketing

Nachdem eine Personalplanung seitens des Unternehmens vorgenommen wurde, kann sich Bedarf hinsichtlich der Rekrutierung von neuem Personal zeigen. Für die Gewinnung dieses Personals eröffnet der Einsatz von KI verschiedene Möglichkeiten:

Zunächst können KI-Tools **Stellenausschreibungen** generieren und personalisieren. Stellenausschreibungen werden mit KI personalisiert, indem Formulierungen gezielt an bestimmte Zielgruppen angepasst werden. KI-Lösungen wie z. B. Textio wurden auf Basis von vergangenen Ausschreibungen und deren Wirkung trainiert, um gezielte Vorschläge zur Verbesserung zu machen und damit die Wirkung bei den Zielgruppen zu erhöhen. Eine solche Optimierung kann bei Stellenausschreibungen erfolgen, die durch die Personalabteilung selbst generiert worden sind. Eine andere Herangehensweise besteht darin, Large Language Models (LLM) zu nutzen, um Stellenausschreibungen komplett automatisiert zu erstellen und sie dann nur noch nachzubearbeiten. Dabei kann KI einen enormen Vorteil ausspielen: Sie kann für verschiedene Recruitingkanäle automatisiert eingesetzt werden und dieselbe Ausschreibung anpassen. Dabei können Anbieter wie z. B. catchtalents den Text, Schlagworte, aber auch die visuelle Gestaltung der Ausschreibung automatisiert adaptieren.

Liegt nun eine Stellenanzeige vor, kann es passieren, dass sich zu wenige Personen mit einem geeigneten Profil hierauf bewerben. Genau an dieser Stelle kann nun KI unterstützen: Unter dem Begriff **Social Recruiting** (auch: Programmatic Job Advertising/Active Sourcing) analysiert KI Social-Media-Profile auf bspw. klassisch arbeitsbezogenen Portalen wie LinkedIn, aber auch auf eher zu privaten Zwecken genutzten Plattformen z. B. Instagram und zeigt gezielt potenziell geeigneten Personen Werbung für die Stellenausschreibung. Zudem können je nach Plattform individualisierte Ansprachen per Direktnachricht erfolgen. Dies erhöht die Chance, mehr Bewerbungen auf eine Stellenanzeige zu erhalten.

Zusätzlich kann es sinnvoll sein, dass beim Besuch einer Stellenausschreibung Bewerber und Bewerberinnen auf der Webseite begleitend die Möglichkeit haben, mit einem Chatbot mehr zur Stelle zu erfahren. Dabei könnten z. B. Informationen zur Unternehmenskultur, detailliertere Informationen zur Stellenausschreibung oder zur Gestaltung der Bewerbung bereitgestellt werden. Dies verbessert erneut das Nutzererlebnis, was die Zahl der Bewerbungen steigern kann.

Schließlich kann KI das Personalmarketing optimieren, indem die Wirkung der Stellenausschreibungen und der Werbung für Stellen systematisch kontrolliert wird. Dies ist wichtig, um zu prüfen, welche Kanäle für Ausschreibungen besonders geeignet und welche eher ungeeignet sind.

## 5.2.3 Personalauswahl

Insbesondere das Thema Personalauswahl und KI ist in den Fokus der Öffentlichkeit und Kritik geraten. Um diese Kritik besser beurteilen zu können, ist es wichtig, die Funktionsweise zu kennen: Basierend auf historischen Daten der aktuellen Belegschaft, vergangenen Bewerbungsdaten und Inhalten der relevanten Stellenausschreibung kann eine KI trainiert werden. Auf Basis dieses Trainings leitet das System Kriterien zur Beurteilung einer Bewer-

bung ab. Um diese einsetzen zu können, müssen in einem zweiten Schritt die eingehenden Bewerbungen maschinenlesbar werden. Dies kann entweder über eine digitale Maske zur Eingabe von Daten oder die originalen Daten einer Bewerbung werden mittels **CV-Parsing** aus den Lebensläufen herausgelesen. Bei letzterem Verfahren können systematische Fehler entstehen, wenn die Maschine mit verschiedenen Lebenslaufformaten nicht gut zurechtkommt und dadurch Informationen gar nicht oder falsch erfasst werden. Hierdurch könnten ggf. geeignete Personen aussortiert werden, weshalb eine Qualitätskontrolle an dieser Stelle essentiell ist. Außerdem ist eine solche Qualitätskontrolle aufgrund des KI-Gesetzes der EU notwendig, denn ein CV-Parsing unterstützt die Personalauswahl, hierbei handelt es sich laut KI-Gesetz der EU um eine Hochrisiko-KI, bei der spezielle Auflagen wie z. B. die Gewährleistung der Datenqualität erfüllt werden müssen, um sie zu nutzen (Gehrman und Förster 2024)<sup>1</sup>. Die erfassten Daten können anschließend in einer Datenbank hinterlegt werden, die sich entweder die Personalverantwortlichen ansehen oder von einer KI mithilfe der zuvor definierten Kriterien vorausgewählt werden. Dabei sollte die KI stets nur eine Vorauswahl treffen, da dann potenzielle Fehler korrigiert werden können. Generell ist bei der Nutzung von solchen datenbasierten Verfahren die DSGVO zu beachten und bspw. eine Zustimmung der Bewerbenden für eine solche Datenverarbeitung einzuholen. Zudem kann die DSGVO dadurch eingehalten werden, dass KI nur die Vorauswahl trifft und ein wesentlich Anteil an der Entscheidung noch beim Menschen liegt (Muti 2024).

Es gibt intensive Debatten in der Öffentlichkeit und in der Forschung zu der Frage, wie gut eine **KI-basierte Vorauswahl** ist. Aufgrund der Nutzung von historischen Daten können Verzerrungen bei der Auswahl,

die bereits in der Vergangenheit passiert sind, sich auch auf die zukünftige Auswahl auswirken, in dem bspw. bestimmte Personengruppen auf Basis von Merkmalen wie Alter oder Geschlecht bevorzugt werden. Zu solchen Verzerrungen ist aber festzuhalten, dass diese einerseits menschliche Fehler, die in den zum Training benutzten Daten enthalten sind, repräsentieren (Goretzko und Israel 2022). Andererseits kann durch ein Trainieren mit qualitätskontrollierten Daten eben solchen Verzerrungen vorgebeugt werden. Zudem sei angemerkt, dass insbesondere auch Personalverantwortliche vielfältige kognitive Verzerrungen besitzen, die genauso zu erheblichen Problemen bei der Auswahl führen können (Kanning 2015). Deshalb kann man nicht einfach sagen, ob Menschen oder KI bessere Ergebnisse bei der Personalauswahl erzielen. Wenn Verzerrungen so gering wie möglich ausfallen sollen, sollten die Personalverantwortlichen möglichst umfangreich methodisch im Bereich Personaldiagnostik ausgebildet und die KI sollte bestmöglich trainiert sein.

Bei der Personalauswahl kann zwischen einem **Exploitations- und einem Explorations-Vorgehen** unterschieden werden (Li et al. 2020). Bei einer Exploitation werden Personen mit einem perfekt passenden Lebenslauf berücksichtigt, wohingegen bei einer Exploration auch solche Bewerbungen berücksichtigt werden, die eher unkonventionelle Merkmale aufweisen, aber trotzdem hochgeeignet für eine Stelle sein können. Insbesondere bei einem Exploitation-Vorgehen besteht das Risiko, dass nur stromlinienförmige Bewerbungen bei der Auswahl berücksichtigt werden und dadurch Potenzial für das Unternehmen verloren geht. Hier zeigen erste Studienergebnisse, dass es sehr entscheidend ist, wie der hinter der KI stehende Algorithmus trainiert wird (Li et al. 2020). Nichtsdestotrotz fehlen Längsschnittstudien, welche die längerfristigen Folgen von KI-basierten Personalentscheidungen für den Erfolg von Unternehmen und die Zufriedenheit des Personals untersuchen. Auch ist es wichtig zu beachten, wie sich der Einsatz von KI auf die Bewerber und Bewerberin-

<sup>1</sup> An dieser Stelle sei auf das ► Kap. 3 von Schwartmann und Köhler in diesem Sammelband verwiesen, das sich umfangreich mit dem KI-Gesetz und dem Thema Datenschutz beschäftigt.

nen auswirkt. Hier konnten bspw. Dahm und Dregger (2019) zeigen, dass insbesondere assistive KI-Systeme<sup>2</sup> gegenüber autonomen KI-Systemen von den Bewerbern und Bewerberinnen präferiert werden. Sofern aber eine KI mit hinreichender Qualitätskontrolle zum Einsatz kommt, kann sie die Auswahl beschleunigen. Ein beschleunigter Personalauswahlprozess kann sich wiederum positiv dahingehend auswirken, dass zum einen das eigene Unternehmen „einen Schritt schneller“ als die Konkurrenz ist und sich zum anderen Bewerber und Bewerberinnen stärker wertgeschätzt fühlen, da sie schneller Rückmeldung zu ihrem Bewerbungsprozess erhalten. Letztendlich sind sowohl KI-basierte als auch menschliche Personalauswahlprozesse nie perfekt, weshalb deren konstante Qualitätskontrolle, egal wie der Prozess gestaltet wird, essentiell ist. Bei einer KI kann die Qualitätskontrolle auf verschiedenen Wegen durchgeführt werden: So müssen Bewerber und Bewerberinnen Feedbackmöglichkeiten erhalten, um z. B. auf Fehler bei der Datenerfassung hinweisen zu können. Darüber hinaus sollten die Ergebnisse der Auswahl regelmäßig z. B. mit statistischen Methoden evaluiert werden, um mögliche Verzerrungen zu erkennen. Eine solche Analyse kann auch eine Prüfung der sogenannten prognostischen Validität umfassen. Prognostische Validität bzw. Vorhersagekraft (Keller et al. 2018) meint in diesem Kontext, ob man statistisch valide z. B. den beruflichen Erfolg einer ausgewählten Person auf Basis der Einschätzung der KI-Vorauswahl voraussagen kann.

Neben einer solchen Analyse von Dokumenten kann KI auch dazu genutzt werden, Verhaltensmuster insbesondere bei videobasierten Jobinterviews einzusetzen. Dabei kann ein Einsatz auf zwei Arten erfolgen: Einerseits können solche Systeme dazu genutzt werden, dass sich Bewerber und Bewerberinnen auf ein Interview vorbereiten, andererseits kön-

nen solche Systeme aber auch real eingesetzt werden und bei der Auswahl von Personal unterstützen. Kern hinter beiden Nutzungsformen ist die Analyse von non-verbalen **Verhaltensdaten**, die z. B. die Blickrichtung, die Mimik oder auch Eigenschaften der Stimme umfassen (Dissanayake et al. 2021). Basierend auf diesen Daten wird ein Muster erkennender Algorithmus trainiert, bei dem am Ende Korrelationen z. B. zwischen non-verbalem Verhalten und z. B. Persönlichkeitsmerkmalen berechnet werden (Dissanayake et al. 2021). Bisher wurden solche Persönlichkeitsmerkmale mittels Persönlichkeitstests bei der Personalauswahl gemessen. Hierbei zeigen sich in Studien eine Vorhersagekraft (prognostische Validität) von z. B.  $r = 0,40$  für den beruflichen Erfolg (Keller et al. 2018). Inwiefern solche Persönlichkeitsmerkmale mithilfe von verbalem und non-verbalem Verhalten durch eine KI erfasst werden kann, wird zurzeit umfangreich in der Forschung evaluiert (Hashemi et al. 2025). Mittels KI wird ein Score für Persönlichkeitsmerkmale anhand z. B. des non-verbalen Verhaltens festgesetzt, der wiederum entweder im Bewerbungstraining dazu genutzt werden kann, Feedback zu geben (Kothari et al. 2024), andererseits aber auch bei der Auswahl von Personal unterstützen kann. Darüber hinaus gibt es Ansätze in der Forschung, solche KI-Systeme auch so zu nutzen, dass sie in der Interviewsystemssituation direkt unterstützen und auf Basis des Zustands des Bewerbenden Vorschläge für die nächsten Fragen im Jobinterview generieren können. Letzterer Ansatz ist aber noch wenig erforscht (Dissanayake et al. 2021). Forschungsergebnisse deuten darauf hin, dass solche Systeme durchaus gute Ergebnisse bei der Erkennung von Persönlichkeitsmerkmalen aufweisen können, jedoch zeigten Dissanayake et al. (2021), dass die Befragten zu 70 % ein klassisches Jobinterview einem KI-basierten Interview vorziehen. Solche Systeme sind von mehreren Anbietern, etwa enablex, auf dem Markt verfügbar, jedoch ist hier schwer zu beurteilen, welche Qualität diese Verfahren haben. Zudem gilt es zu bedenken, dass der Artificial

2 Assistiv meint an dieser Stelle, dass die KI nur unterstützend in den Entscheidungsprozess eingreift, nicht aber eigenständig, also autonom Entscheidungen trifft (Kreipp 2021).

Intelligence Act der Europäischen Union (EU AI Act), der am 01.08.2024 in Kraft getreten ist, solche System als Hochrisiko-KI einstuft und deshalb besondere Auflagen zu beachten sind (Kock 2024). Wie bereits zuvor erwähnt, gelten bei solchen Hochrisiko-KI besondere Bedingungen z. B. in Bezug auf die Datenqualität.<sup>3</sup>

## 5

### 5.2.4 Onboarding von Personal

Ein letzter Baustein des Recruiting stellt die Nutzung von Sprachassistentensystemen im Rahmen des Onboardings dar. Neu eingestelltes Personal hat zu Beginn viele Fragen zu Prozessen und Verantwortlichkeiten innerhalb einer Organisation. Dies kann dazu führen, dass z. B. Führungskräfte immer wieder mit Fragen überhäuft oder alternativ bestimmte Fragen von eben jenem neuen Personal nicht gestellt werden, um Belastungen der Führungskräfte zu vermeiden. Genau hier können mittlerweile recht zuverlässig und umfangreich LLM-basierte Sprachassistentensysteme zum Einsatz kommen. Entscheidend für ihre Qualität ist zunächst das Vorhandensein eines qualitativ hochwertigen Datensatzes mit Informationen z. B. zu Abläufen, etwa wie in einem Unternehmen ein Urlaubsantrag zu stellen ist. Mithilfe dieser Informationen können dann sogenannte **Retrieval-Augmented-Generation-(RAG-) Chatbots** zuverlässige Antworten generieren (Lewis et al. 2021). Die Idee hinter RAG-Chatbots besteht darin, dass LLM-basierte Chatbots wie z. B. ChatGPT zwar über ein sehr gutes Allgemeinwissen verfügen, ihnen aber das konkrete und aktuelle Wissen zum Unternehmen fehlt. Anstatt nun mit hohen Kosten das LLM nachzutrainieren, ruft das System ergänzend zunächst Informationen des Unternehmens wie z. B. eine Prozessdokumentation zum Thema *Urlaubsantrag* ab und erzeugt dann mit den daraus gewonnenen Informationen eine Antwort. Zur Optimierung der Antwortformulierung unterstützt dann ein

LLM. Der Vorteil eines solchen Systems liegt nicht nur in den geringen Kosten im Vergleich zum Nachtrainieren eines LLM, sondern auch darin, dass sensible Informationen zu Prozessen im Unternehmen nicht an Externe gehen. Zusätzlich wird verhindert, dass durch sogenannte „Halluzinationen“, die bei LLM-basierten Systemen auftreten können, falsche Informationen herausgegeben werden.

Halluzinationen bedeutet hierbei, dass eine KI Informationen als Fakt ausgibt, die irreführend oder objektiv falsch sind (Smith 2023). Ursächlich hierfür können verschiedene Faktoren wie z. B. die Daten oder auch die Konstruktion der generativen Modelle sein. Am Begriff „Halluzinationen“ gibt es aber auch Kritik (Maleki et al. 2024).

Damit ein solches System gut von der Belegschaft angenommen wird, ist neben der richtigen IT-Architektur eine gute Gestaltung des Gesprächs zwischen dem Chatbot und dem Nutzer oder der Nutzerin mithilfe von **Conversational Design** nötig. Hierbei geht es darum zu klären, über welche Funktionalität ein solches System verfügen und wie es mit Nutzer und Nutzerinnen interagieren soll. Hierbei kommt es auf die Gesprächsführung an, die von zentralen Aspekten wie einer künstlichen Persönlichkeit des Systems beeinflusst werden kann (Dregger 2023). Künstliche Persönlichkeit meint hierbei, dass ein Mensch einer Maschine Persönlichkeitsmerkmale aufgrund ihres Verhaltens zuschreibt, etwa dadurch, wie sie spricht. So kann z. B. eine Maschine als mehr oder weniger verträglich wahrgenommen werden (Holtgraves et al. 2007), was einen Einfluss auf die Nutzerakzeptanz hat. Die Art der Gesprächsführung, die Gestaltung der Sprache sowie die Qualität der Antworten und Funktionen entscheidet dann darüber, ob ein System in einer Organisation akzeptiert und tatsächlich genutzt wird. Beispielsweise ist es wichtig, wie die Begrüßungsnachricht eines Chabots gestaltet ist. Diese muss rechtlichen Anforderungen genügen, aber gleich-

3 Mehr hierzu siehe ▶ Kap. 3 in diesem Band.

zeitig auch einen natürlichen Gesprächsfluss eröffnen (Van Hooijdonk et al. 2023).

## 5.2.5 Personalentwicklung

Damit die Qualifikationen der Beschäftigten mit sich ändernden Anforderungen im Einklang bleiben, ist deren stetige Weiterqualifizierung wesentlich. Auch hierbei kann KI auf vielfältige Art unterstützen. Der folgende Abschnitt umreißt hierfür die Themen Wissensmanagement, personalisierte Fort- und Weiterbildung und Gesundheitsmanagement.

### ■■ Wissensmanagement

Durch den Abgang der Babyboomer-Generation im Zuge des demographischen Wandels droht Organisationen, dass wichtiges Wissen verloren geht und vakante Stellen unbesetzt bleiben (Freudung et al. 2023). Auch hier kann KI helfen: Einerseits kann sie dazu dienen, Prozesse effizienter zu gestalten und somit Lücken, die durch das Abtreten der Babyboomer-Generation entstehen, zu kompensieren. Andererseits kann KI bei der Erfassung von Bestandswissen unterstützen: So können z. B. Interviews mit der Babyboomer-Belegschaft zu den Prozessen in den Unternehmen geführt werden, um insbesondere an verborgenes bzw. implizites Wissen zu gelangen. Mithilfe von LLM-Methoden können die wesentlichen Aussagen dann zusammengefasst und für die weitere Verwendung aufbereitet werden. Beispielsweise könnten Prozessbeschreibungen automatisiert durch die Ergebnisse der Interviews verbessert werden. Außerdem kann das Wissen dazu genutzt werden, die Informationen mithilfe von Sprachassistenten wie Chat- oder Voicebots der Belegschaft bereitzustellen.

### ■■ Fort- und Weiterbildung

Fort- und Weiterbildung sind zentrale Elemente, um Wissenslücken bei Beschäftigten einer Organisation zu schließen und neues Know-how aufzubauen. Dabei müssen die Inhalte nicht nur zur Organisation, sondern auch

zur Belegschaft passen bzw. deren individuellen Bedürfnissen und Interessen Rechnung tragen. Hierbei kann KI zunächst helfen, die „richtigen“ Personen in einer Organisation für eine bestimmte Weiterbildungsmaßnahme zu identifizieren. Dies setzt voraus, dass aktuelle Daten zur Qualifizierung der Beschäftigten vorliegen. Zudem sollten auch individuelle Wünsche des Personals als Datenpunkt in den Prozess einfließen. Darüber hinaus kann KI ihre Vorteile bei digitalen Lernformaten ausspielen. Insbesondere Methoden der **generativen KI** können bspw. helfen, Lernmaterialien zu erstellen und diese adaptiv auf die Bedürfnisse der Belegschaft zuzuschneiden (Fischer et al. 2024). Präferiert eine Person z. B. eher kleinere Lerneinheiten, so kann der Inhalt auf solche heruntergebrochen werden. Eine weitere Möglichkeit der individuellen Unterstützung durch KI kann in einem Chatbot-System bestehen, mit dem sich interessierte Mitarbeiter und Mitarbeiterinnen zu Weiterbildungsmöglichkeiten informieren können.

### ■■ Gesundheitsmanagement

Schließlich kann auch die Gesundheit der Beschäftigten in Organisationen gezielt durch KI gefördert werden. Dies umfasst sowohl Coaching- und Weiterbildungsmaßnahmen als auch Monitoring-Methoden, die individuelles **Gesundheitsfeedback** ermöglichen. In Bezug auf klassische Weiterbildung und Coaching existieren inzwischen zahlreiche Anbieter, die Mitarbeitenden ermöglichen, sich durch Apps zum Thema Gesundheit weiterzubilden. Solche Apps nutzen digitale Inhalte, die interaktiv mithilfe von zuvor beschriebenen KI-Methoden aufbereitet werden. Darüber hinaus kann mithilfe von KI der Gesundheitszustand der Beschäftigten eines Unternehmens erfasst und damit auch verbessert werden: Beispielsweise bietet das Unternehmen Corvolution GmbH eine KI-basierte Lösung namens *mesana* hierfür an. Zunächst werden hierzu – ausschließlich mit Zustimmung der Mitarbeitenden – Daten wie bspw. die subjektive Wahrnehmung des eigenen Gesundheitsverhaltens oder der Belastung erhoben. Dies erfolgt mittels Fragebögen.